

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

A Thesis Submitted for the Degree of PhD at the University of Warwick

<http://go.warwick.ac.uk/wrap/36380>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

**Programme evaluation by teachers: an
observational study**

by

Richard Kiely

A thesis submitted in partial fulfilment of the
requirements for the degree of Doctor of Philosophy

University of Warwick
Centre for English Language Teacher Education

March 2000

Table of contents

	Page
Table of contents	ii
List of figures	vii
List of Appendices	x
Acknowledgements	xi
Declaration of published work	xii
Abstract	xiii
Abbreviations	xiv
Data Conventions	xv
Chapter 1 Introduction	1
Chapter 2 Discourses of Programme Evaluation in English Language Teaching	10
2.1 Introduction – a conceptual map	10
2.2 Evaluation - Three phases	12
2.3 Domains of language programme evaluation	15
2.4 Evaluation in ELT - Two research traditions	18
2.5 Programme theory in ELT	19
2.6 The Bangalore evaluation and the development of ELT evaluation theory	24
a) The role of experimental design	26
b) The measurement of outcomes	29
c) The need to describe component parts of the programme	31
d) The need to involve programme participants in the programme	36
2.7 Programme evaluation by teachers	38
2.8 Other traditions of teacher inquiry	40

	Page
a) Action research	41
b) Reflective practice	43
c) Classroom ethnography	44
d) Teacher evaluation in the appraisal tradition	45
2.9 Teacher Inquiry - Summary	47
2.10 Summary	48
Chapter 3 Pilot study - teachers evaluating an EAP programme	52
3.1 Introduction	52
3.2 Institutional context	52
3.3 The programme	56
3.4 Research methodology	57
3.5 The evaluation of the programme	59
a) Stage 1 - Baseline evaluation - establishing the pedagogy	59
b) Stage 2 - Formative evaluation - taking stock and moving on	61
c) Stage 3 - Summative evaluation - preparing for next time	68
3.6 Lessons from the Pilot Study	70
a) Wider implications	70
b) The study of evaluation - methodological implications	70
c) The study of evaluation - conceptual foci	72
3.7 Summary	76
Chapter 4 Research design and methodology	77
4.1 Introduction	77

	Page
4.2	Evaluation and Research 77
4.3	Research approach 1 - Issues of ontology and epistemology 79
4.4	Research approach 2 - Case study 84
4.5	Research approach 3 - Ethnography 88
4.6	Research strategy 1 - Procedures and instruments 90
a)	Classroom observation 91
b)	Interviews 93
c)	Recording of one group of students holding an evaluation discussion 95
d)	Documentation 96
4.7	Research strategy 2 - Data analysis and interpretation 97
4.8	Validity as a research issue 101
4.9	Validity as an evaluation issue 104
4.10	Summary of research design 107
4.11	Summary 109
Chapter 5	Programme constructs and evaluation processes 110
5.1	Introduction 110
5.2	The programme - three perspectives 111
a)	The specification of the institution 111
b)	The construct of the teacher 112
c)	The series of classroom activities 113
d)	Constructs in conflict 116
5.3	The formal evaluation event 117
a)	The evaluation 117
b)	Commitment of the teacher 122
c)	Language skills development 124
d)	Equal status of all voices 125

		Page
	e) The role of the teacher	132
5.4	The values underpinning the programme	136
5.5	Some consequences of the evaluation	143
	a) Characterisations of vocabulary	144
	b) Programme improvement and teacher development	149
5.6	Summary	153
Chapter 6	The role of the institution - sleeping partner or remote control?	157
6.1	Introduction	157
6.2	The institution in the programme - issues of positioning and representation	158
	a) Boundary management	160
	b) Role and status of teachers	162
	c) Role of students	164
6.3	The programme in the institution - the inevitably gapped loop	166
	a) The customising of NGT	166
	b) Evaluation and programme development	168
	c) Evaluation and innovation - long and short cycles	170
6.4	The wider agenda - evaluation, quality, and the framing of the teacher-manager relationship	174
	a) The evaluation policy	175
	b) Quality management in higher education	179
6.5	Summary	184
Chapter 7	'In fact I can't really lose' - the case of Laure	185
7.1	Introduction	185
	a) Why Laure?	187

		Page
7.2	The social dimension of learning in the programme	189
7.3	The materials: Laure's likes and dislikes	192
7.4	Negotiating an essay title	199
7.5	A note of paradigm and method	205
7.6	Summary	210
Chapter 8	Conclusion: Rethinking, repositioning and rewiring evaluation	211
8.1	Introduction	211
8.2	The different traditions of programme enquiry	213
8.3	Institutional policy and evaluation design	216
8.4	An ethical framework	219
8.5	Evaluation and teaching	221
8.6	Programme evaluation: towards an integrated management strategy	223
8.7	Limitations of this study	225
8.7	Concluding remarks	226
	Endnotes	227
	Bibliography	231
	Appendices	247

List of Figures

Figures		Page
	Chapter 2	
	Discourses of Programme Evaluation in English Language Teaching	
2.1	Conceptual map of Chapter 2	11
2.2	Summary of three conceptions of valuation	14
2.3	Four evaluation accounts of ELT programme theory	23
2.4	Features of different teacher inquiry traditions	48
2.5	Multi-functional evaluation	49
	Chapter 3	
	Pilot study - teachers evaluating an EAP programme	
3.1	Departmental approach to evaluation of programmes	54
3.2	Group discussion stage of NGT – CS1:5	62
3.3	Plenary discussion stage of NGT – CS2:5	64
	Chapter 4	
	Research design and methodology	
4.1	Assumptions underlying inquiry in the social sciences (after Cohen and Manion 1994:9)	80
4.2	Naturalism and positivism (after Lincoln and Guba 1985:37)	82
4.3	Summary of values underpinning research design	83
4.4	Features of case study approach related to this study	86
4.5	Reasons for using ethnography as form of naturalistic inquiry	89

		Page
4.6	Summary of data collected on two programmes (CS1 & CS2)	90
4.7	Strategies for analysis of ethnographic data (after Goetz and LeCompte 1981, Lincoln and Guba 1985)	99
4.8	Techniques for establishing trustworthiness	102
4.9	Techniques for development of researcher role	103
4.10	Determining evaluation validity (from Messick 1989:20)	105
4.11	Summary of research design for main study	107
 Chapter 5		
Programme constructs and evaluation processes		
5.1	Construct of the programme as represented in the module guide	111
5.2	Construct of the module as represented by the teacher	112
5.3	Programme materials - weeks 1-12	115
5.4	Programme materials - weeks 1-5 (pre-evaluation event)	115
5.5	Programme materials - weeks 6-12 (post-evaluation event)	115
5.6	Profile of the participants in the audio recorded small group discussion	118
5.7	Content analysis of the recorded small group discussion	119
5.8	Content analysis of the plenary feedback session (CS2:5)	120
5.9	Anna's response to evaluation feedback on vocabulary encountered in texts	150

	Page
Chapter 6	
The role of the institution - sleeping partner or remote control?	
6.1	Summary of the off-campus examination episode 158
6.2	Stakeholder alliances 165
6.3	Comparison of recommended procedure and actual evaluation 167
	Models of universities as organisations 182
Chapter 7	
'In fact I can't really lose' - the case of Laure	
7.1	Schedule for coursework assessment 199
Chapter 8	
Conclusion: Rethinking, repositioning and rewiring evaluation	
8.1	Summary of research foci, key findings, and further exploration 213
8.2	Examples of different programme inquiry traditions in action 215
8.3	Summary of evaluation policy development 217
8.4	Framework for evaluation development 225

	List of Appendices	Page
Appendix 1	Data used for this study	247
Appendix 2	Questionnaire on Evaluation for teachers of EAP	248
Appendix 3	Interview schedule 1 - Teacher	252
	Interview schedule 2 - Students	253
	Interview schedule 3 - Manager	254
Appendix 4	CS2 – Interviews with teacher, coordinator and students	255
Appendix 5	CS2 – Classroom materials	256
Appendix 6	Evaluation policy of the institution	259
	Evaluation policy of the department	263
Appendix 7	Writing task CS2:1	264
	Writing task CS2:5	266

Acknowledgements

Many people have helped me in completing this thesis. I would like to express thanks to:

Dr. Pauline Rea-Dickins, who as a colleague and friend over many years as well as supervisor, made it all possible;

Anna and Millie, the teachers who gave me time and access to their classrooms and pedagogy;

the different professional communities I have been fortunate to work with, especially Applied Linguistics colleagues at Thames Valley University, University College Chichester, and the University of Warwick, and the EAP students who so willingly shared their experience with me;

Linda, Ellen and Laura who, in so many different ways, encouraged me.

Declaration of published work

A version of Chapter 3 has been published as:

Kiely, R. 1998 Programme evaluation by teachers: issues of policy and practice. in Rea-Dickins, P. and K. Germaine 1998 (Eds.) *Managing Evaluation and Innovation in Language Teaching: Building Bridges*. Harlow: Addison Wesley Longman pp. 78-104

A version of Chapter 5, Section 5.5 has been published as:

Kiely, R. 1999 Classroom evaluation: a platform for teacher development. in J. Edge & K. Richards (Eds.) 1999 *Proceedings of Teachers Develop Teachers' Research Conference (TDTR4)* Leuven Belgium. CD - ROM

The data collected and analysed for this study has generated the following articles:

Kiely, R. 1999a Evaluation by teachers. *IATEFL Issues* 147 pp. 14-16

Kiely, R. 1999b Taking stock: mid-course evaluation. *IATEFL Testing, Evaluation and Assessment SIG Newsletter* pp. 6-9

Abstract

Evaluation is a term used to denote many different processes in Applied Linguistics and language education, from language proficiency assessment to materials selection to project management. In this study evaluation is taken as programme inquiry for the purposes of accountability and development (Rea-Dickins and Germaine 1992; Weir and Roberts 1994). The context of evaluation is EAP programmes in higher education, where a policy of programme evaluation by teachers provides base data for the institution's quality management system.

Evaluation in this sense is a complex process, embedded in and potentially destabilising for the prevailing social order in classrooms, departments, and institutions. To capture this complexity and embeddedness, a qualitative methodology is developed from naturalistic inquiry and ethnography to study two EAP programmes and their evaluations. The classroom observation and interview data are used to explore the operation of the evaluation policy, its consequences in terms of programme improvement and teacher development, its quality assurance functions, and its impact on the programme experience of one particular student.

The study shows that in this context the evaluation policy has beneficial and problematic consequences. It looks good, and thus enables a case to be made for rigorous quality management. It provides a space for students to negotiate the programme, and for the teacher, opportunities for reflection and development. However, the policy also risks being seen by teachers as intrusion, compromising teacher autonomy, and aggravating teacher-manager problems. Students provide feedback on teaching and inputs (including the teacher) rather than on learning or opportunities for learning in terms of the values underpinning the programme. The strong students can use the evaluation strategically, to further their own preferences, at the expense of the weaker students who are marginalised.

The concluding chapter outlines some ways forward for this approach to evaluation in this context. These include realigning purposes and methods, building on teacher ownership of and responsibility for programmes, linking evaluation in the classroom to other forms of educational inquiry, and managing the programme in its widest sense in an ethical framework.

Key words

English Language Teaching, Programme evaluation, Nominal Group Technique, Classroom ethnography, Teacher development, Quality management, Learning strategies

Abbreviations

BALEAP	British Association of Lecturers in English for Academic Purposes
BASELT	British Association for State English Language Teaching
BEP	Bilingual Education Project
CARN	Classroom Action Research Network
CLT	Communicative Language Teaching
CS1	Case Study 1
CS2	Case Study 2
CTP	Communicational Teaching Project
DfID	Department for International Development
EAP	English for Academic Purposes
ELT	English Language Teaching
ELTJ	English Language Teaching Journal
ERASMUS	European student exchange programme
ESP	English for Specific Purposes
HEFC	Higher Education Funding Councils
HEFCE	Higher Education Funding Council for England
MBA	Master of Business Administration
NGT	Nominal Group Technique
ODA	Overseas Development Administration
OFSTED	Office for Standards in Education
QAA	Quality Assurance Agency
SA	Self Access
SAD	Self Assessment Document
SLA	Second Language Acquisition
TBL	Task-based Learning
TQA	Teaching Quality Assessment

Data Conventions

- i) To preserve the anonymity of informants, the names of all teachers and student have been changed. In addition, the teachers are referred as female, regardless of real gender.
- ii) The two EAP programmes studied are referred to as Case Study 1 (CS1) and Case Study 2 (CS2).
- iii) Fieldnotes (classroom observation notes augmented by headnotes) are referred to throughout as **classroom notes**. The actual words of speakers are in unmarked font, while my observations are in italics.
- iv) The interviewer (Int.) in the interview extracts quoted is, in all cases, myself.
- v) The transcription of spoken English in the classroom notes and interview extracts is designed to facilitate easy reading, especially by teacher and student informants. To achieve this goal, I have punctuated the text and corrected some minor infelicities. I hope however, I have maintained the *voice* and *character* of the original.
- vi) The code for data in the text (e.g. CS2:5) refers to the programme (CS2) and the session / week (Week 5)
- vii) The code for classroom materials (e.g. 5/3/3) refers to the date, month, and place in the series of materials used in that session.

Chapter 1

Introduction

This study is about programme evaluation as carried out by English Language teachers in their own classrooms. The focus on this particular topic is determined by my own professional experience. As a teacher, programme coordinator, and project leader in the UK higher education sector, I became aware of a tension between the theory and practice of evaluation; between the literature on the properties and potential on the one hand, and the uncertainties and discontinuities which characterise practice on the other. The central thesis of this study is that the potential of evaluation is equally the source of the problems with practice: the use of evaluation for both pedagogic and management purposes - improvement to teaching and learning on the one hand, and for quality assurance and administrative purposes on the other - runs the risk of compromising both tasks. This thesis is explored:

- i) through analyses of the discourses of evaluation, in English Language Teaching (henceforth ELT), Applied Linguistics and in educational and social programmes more generally, as represented in the literature; and
- ii) through the empirical study of two English for Academic Purposes (henceforth EAP) programmes where evaluation, as set out in institutional and departmental policies, is carried out by teachers.

Over recent decades there has been an increasing interest in evaluation in ELT, corresponding to a similar trend across the whole range of educational programmes, and indeed social programmes and services in general. In the public sector, led by initiatives in the United States and Britain, there has been a sustained effort to bring transparency and accountability to activities funded from the public purse. A central goal here has been efficiency - in policy terms, a Thatcher-Reagan type slimming

Chapter 1

down of the public sector generally by imposition of features of the market economy. In the private sector, providers of services have been motivated to understand better the needs and wants of their clients. The goal here has been effectiveness - rendering the service provided more finely-tuned to the particular needs of clients and user-groups. Together these twin concerns - efficiency and effectiveness - have come to characterise quality. The management task in providing services in the health, social services, and training fields as well as in the general education and ELT fields, is to evaluate all aspects of the programme in question, so as to a) continuously improve it, and b) be able to demonstrate to those outside the day to day operations that there is appropriate regard for the exigencies of efficiency and effectiveness.

This trend has had implications for characterisations of evaluation. In the 1970's evaluation focussed on policy formulation and accountability:

Educational evaluation is the process of delineating, obtaining and providing useful information for judging decision alternatives.

Stufflebeam *et al* 1971: 43

Systematic educational evaluation consists of a formal assessment of the worth of educational phenomena.

Popham 1975

By the 1990's it had become a means of relating policy to practice and vice versa, and of understanding how these are constantly changing. In the words of Norris:

Evaluation has emerged as the major practical expression of the application of theories and methods from applied social science to the problems posed by piecemeal social engineering.

Norris 1990: 1

Mitchell illustrates how evaluation as a form of inquiry into whole programmes might complement conventional educational research:

Chapter 1

Evaluation is about inferring the most likely relationships among a whole network of complex events, from a wide range of quantitative and qualitative evidence, not about determining strong causal relationships between small numbers of events. It is about the monitoring of intended events and their effects, but also about the identification of the unexpected, and the proposal of untried solutions.

Mitchell 1990:11

Rea-Dickins elaborates on the potential links between evaluation, quality and programme management:

In addition to focussing on the evaluation of quality processes in language teaching, there will be greater appreciation of the different role(s) of the evaluator(s), both internal and external, and of the relationships that hold between practitioners, sponsors and other relevant stakeholders. If evaluation in ELT is to be effective, we will see a stronger integration of evaluation within practice, as part of an individual's professionalism and an increase in collaborative activity where teachers (and other relevant participants) are actively engaged in the monitoring process.

Rea-Dickins 1994:84

Pennington links 'monitoring' to 'understanding' in the context of language programmes:

Language programme evaluation is then, less a set of figures or documents than it is a set of activities. These activities involve people and their interaction in gaining increased understandings which allow them to function more effectively in their work environment. At the same time as these activities make it possible for people to adapt to their environment, they also open up the way for changing the environment, so that it better suits their needs and purposes. Thus, evaluation at its most basic level is the process of interaction that dynamically relates to people, processes and things that make up a language programme in a process of mutual enlightenment, adaptation, and betterment.

Pennington 1998:205

Legutke and Thomas shows how evaluation activity in the classroom is a requirement for learning within Communicative Language Teaching pedagogy:

Engaging learners in communicative encounters, especially if their aim is to explore emotional content and experiences, can become too bound up in itself unless this activity also reaches an evaluation stage. Trying to understand what has happened while undertaking a particular task, why it was suggested by the teacher, and contributing actively to the evaluation of learning arrangements, sequences, resources and input materials by means of reflection and meta-communicative discourse - all these are considered indispensable learner activities in ELT.

Legutke and Thomas 1991:65

These glimpses of different discourses of evaluation illustrate its potential for the development of individual programmes and learning experiences on the one hand, and the wider field of ELT on the other. However, practice is equally determined by negative elements of evaluation, for example, in how it recasts the role of the teacher in language programmes. First, as evaluation within the programme shifts from a focus of assessment of learning outcomes, or testing, to some form of judgement of the teaching-learning process as a whole, the teacher's performance rather than that of the students, becomes the focus of attention. Second, it commits the teacher to engaging with students as part of the process of programme design, and selection of materials and teaching activities. Thus, the teacher's control over the performance to be judged is diluted. Third, the interactions involved in this are documented to allow the institution to assure quality, which may mean a transfer of ownership on the programme from the teacher to the institution. Parton (writing about evaluation practices in the field of social work) illustrates how the monitoring implicit in evaluation constitutes a shift of ownership:

The increased emphasis on management, evaluation, monitoring and constraining professionals to write things down, is itself a form of government of them, [...]. It forces them to think about what they are doing and hence makes them accountable against certain norms. In the process, power flows to the centre or agent who determines the professionals' inscriptions, accumulates them analyses them in their aggregate form, and can compare and evaluate the activities of others who are entries in the chart.

Parton 1994:26

Parton's view has resonances for current trends in British education, where agencies such as the Office for Standards in Education (OFSTED), the Higher Education Funding Councils (HEFCs), and the Quality Assurance Agency (QAA) determine how their legislated responsibilities can be discharged. The procedures of these agencies inform policy within educational institutions, which in turn shapes, and improves, practice within classrooms. The question is whether this role for evaluation in educational management realises the potential suggested by Mitchell, Rea-Dickins, and Pennington (op.cit.) or leads to the compromised professionalism suggested by Parton, and the attendant conflicts of interest and rifts between policy and practice.

This study takes evaluation as social action, both reflecting ideologies and ontological axioms, and creating cultures and discourses within communities of practice. It takes an empathetic perspective to inherently conflicting positions, seeking to reveal both the theoretical coherence of each, and the ultimate incoherence which emerges when they are integrated in policy and practice. The research strategy reflects the search for cultural understanding, and develops a reflective, ethnographic approach which revisits issues of research design as an integral part of the data analysis and interpretation process. Thus, the research questions guiding the study of evaluation are set out initially in Chapter 2, refined as a result of the pilot study in Chapter 3, and operationalised further in Chapter 4 - the research methodology chapter - in relation to data construction and analysis processes relevant to the research foci and questions.

Chapter 2 of this study surveys the discourses which have informed evaluation as a curriculum development and management tool in ELT in recent decades. It examines how the 'phases' in evaluation theory development in the social sciences generally have been interpreted in the different 'domains' of evaluation practice - mainstream education programmes, foreign language programmes based on SLA traditions of research, and ELT projects. Through a focus on the evaluation of the Communicational Teaching Project - the Bangalore project - the evolution of methodological issues in the 1980s and 1990s is traced. Key factors here - the

description of programme processes and the involvement of programme participants - are linked to the development of qualitative approaches. Finally, the traditions of inquiry used by teachers - action research, reflective practice, classroom ethnography and teacher appraisal - are examined, to determine overlaps and distinctions in this diversity of concepts and labels.

Chapter 3 reports on a pilot study which explores, on the one hand, the extent to which the discordant discourses outlined above are evident in the evaluation of one specific programme, and on the other, how a qualitative research strategy appropriate for the issues and concepts focussed on might be elaborated. The study sets out the institutional and programme contexts, and documents the process of evaluation - baseline at the outset, formative in the fifth week, and summative in the twelfth and final week. Finally, a range of issues are considered: the methodological implications of this pilot for the main study, the merging of evaluation and pedagogy, the students' perception of their role, and the contribution of the institutional policy. This discussion of the pilot study serves to refine the research questions outlined in the preceding chapter.

Chapter 4 develops in detail the research strategy for the main study. The differences between research and evaluation as traditions of inquiry are explored, and the paradigm debate common to both, is discussed to elaborate a methodology which in this case matches ontological and epistemological positions and research focus and purpose. The resulting research approach emphasises the local, the particular, the coherence of cultural practices forged by conflicts and interests within communities of practice. Each case - programme, learning activity, teacher, student - represents itself, explores links between abstracted issues, and provides resonances for readers to consider in relation to their own circumstances. The research design is informed by naturalistic inquiry, and the processes of data collection and analysis by ethnography. These approaches to inquiry underpin the validity of the research, both in terms of data and interpretations.

Chapter 5 explores the ways in which the programme is the evaluand. Exploring the nature of the programme three constructs are set out: the institutional specification from the module guide, the teacher's concept from interview and classroom observation data, and the classroom process, as determined by analysis of the materials and activities of the complete 36-hour programme. The evaluation event is then examined in relation to three programme constructs. This involves exploration of issues which, from the data, seem to shape the evaluation. A detailed analysis of one classroom activity illustrates how it informs evaluations of the programme, both formally in the evaluation event, and informally as in the interview data. Finally, there is a discussion of the consequences of the evaluation, both in terms of changes to the programme, and changes in teacher thinking. The latter permits some discussion on how this consequence might illustrate some ways in which evaluation links to teacher development.

Chapter 6 engages with another consequence of the evaluation: the extent to which it discharges its quality assurance function for the institution as a whole. The focus on the institution is realised first through an analysis of the representations of the institution within the programme, i.e. in the classroom. Then the relationship between the programme and the department within which it is located - the unit with formal responsibility for quality assurance - is analysed. The issues raised lead to quality management debates in the wider higher education sector, and there is tentative linking of factors observed in the evaluation of the programme in question, and perspectives from this wider debate. The issue is the extent to which policy and practice in relation to evaluation and quality management constitute a shared agenda, where different stakeholders seek collaboratively to develop teaching and learning, and conversely, the extent to which there is no shared agenda, and positions taken have a strategic purpose in an on-going competition between interests.

Chapter 7 looks at the experience of one student, Laure, within the programme. The data provide a detailed account of her participation both in the programme and in the evaluation event. The discussion of what Laure expects from the programme, and her response to a range of materials and learning activities problematises the issue of students as programme evaluators. The analysis of how Laure negotiates an essay title suggests on the one hand a strategy of avoidance, together with poor learning strategies. Her positive evaluation of the programme is seen as linked to her ultimate success in it - an over-arching strategy of being positive and engaging may prevent her from becoming discouraged through an appreciation of her English language problems in terms of the programme objectives. Drawing such conclusions, even tentatively, is discussed as a research problem, which can be best resolved through analysis of data from a theoretical ethnographic perspective.

Chapter 8 draws some tentative conclusions from the study, and explores the lessons which might be extrapolated for policy makers and practitioners in evaluation contexts such as that described in this study. The lessons are not set out as generalisations; rather principles which readers might see as resonating with aspects of their situation, and capable of illuminating the way forward. The lessons relate to five aspects of evaluation theory, policy and practice informed by both the library and empirical elements of this study:

- The different traditions of programme inquiry
- Institutional policy and evaluation design
- An ethical framework
- Evaluation and teaching
- Programme evaluation: towards an integrated management strategy.

This study as a whole holds evaluation - the topic of the study - and research - the method of understanding - as separate entities. This dichotomy is both for facility of inquiry, and for exploration in a metaphorical sense of the potential of programme evaluation. The reader is invited thus to consider three questions: first, whether

evaluation by teachers in ELT, as currently constructed, is comprehensively represented; second, whether the method of inquiry is appropriate; and third, whether the merging of topic and method in this study establishes a credible argument for rethinking this form of evaluation such that it may achieve its potential.

Chapter 2

Discourses of Programme Evaluation in English Language Teaching

2.1 Introduction - a conceptual map

This chapter examines the development of evaluation theory in the field of English Language Teaching. It looks at the diverse purposes of evaluation, and the domains in which evaluation theory and practice have been developed: social programmes generally, language programmes, and specifically English language teaching programmes. The impact of different research traditions on evaluation theory and practice, the role of evaluation in the development of an integrated programme theory for ELT, and some problematic issues in the conception of programme in the ELT field are explored. The detail of evaluation theory and practice development is examined through a critical review of a well-known evaluation in recent decades - the Davies-Beretta evaluation of the Bangalore Communicational Teaching Project. Four central issues - the role of experimental design; the measurement of outcomes; the need to describe component parts of the programme; and the need to involve participants in the evaluation - are examined, both in terms of what was done in the Bangalore evaluation and what was done in subsequent evaluations. The Bangalore evaluation is seen as a stimulus in the field of ELT for the development of evaluation. It is appropriate then that the main developments in evaluation through the 80's and 90's are located in a discussion of the key issues in the Bangalore evaluation.

This literature is important in that it frames, conceptually and methodologically, the specific form of evaluation studied in this thesis: programme evaluation by teachers. Other traditions of classroom inquiry - action research, reflective practice and classroom ethnography - and teacher evaluation within the teacher appraisal tradition which contribute to the detail within this evaluation frame are discussed. These traditions are relevant because they inform and develop the language curriculum in ways similar to evaluation. The task in this chapter is to draw on the wider discourses

of evaluation to develop a better understanding of evaluations carried out by teachers within their own programmes. This in turn identifies issues which will be the focus of the empirical parts of this study. Figure 2.1 represents a summary of the conceptual structure of this chapter.

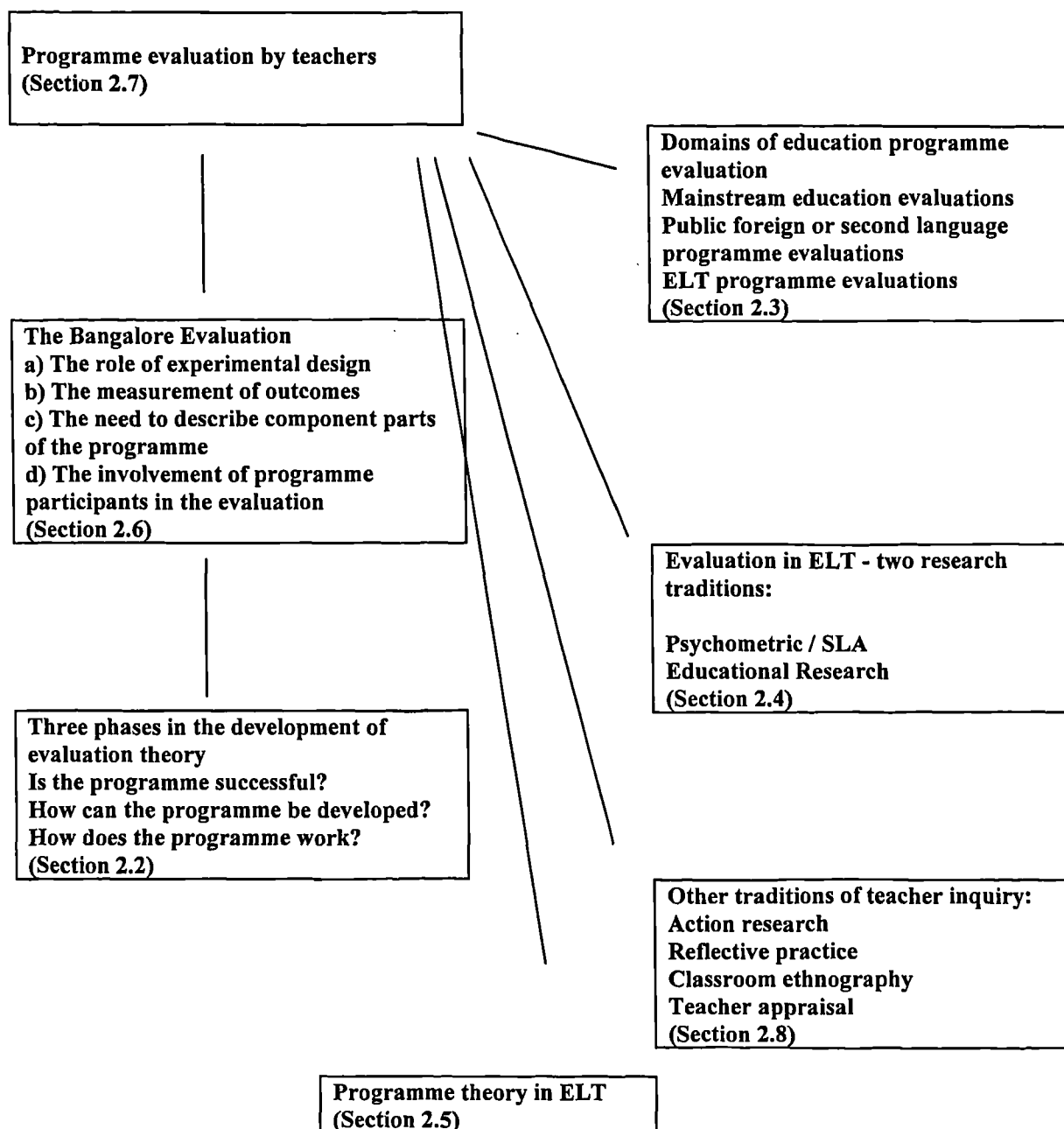


Figure 2.1 Conceptual Map of Chapter 2

2.2 Evaluation - Three phases

Developments in the evaluation of social programmes can be represented in three phases (see Figure 2.2). In some ways these phases represent historical developments, and are often represented in a chronological sequence - Brown (1989); Beretta (1992a); Lynch (1996); Pawson and Tilley (1997). Here, the phasing is seen as a gradual improvement of the theory through focussing on the perceived weaknesses of the “preceding” approach. The phases, or conceptions, also represent a paradigm debate, where the purposes and underlying epistemologies in evaluation are not areas of consensus, but part of an ongoing search for a theory and practice which meets the needs of all evaluation stakeholders (see Chapter 4 below).

The first phase views evaluation as a process of determining whether a given programme, or a specific innovation within that, is successful or not. Answering this question contributes to policy development within the programme area. The research designs combine a positivist epistemology and experimental methodology. The potential impact of the innovation to be evaluated is considered wholly knowable and reducible to a set of measurable criteria. Thus actual impact can be determined through comparison of an experimental group which experienced the innovation, and a control group which did not. Beretta (1992a) summarises a range of evaluations designed and carried out along these principles in foreign or second language programmes, and perhaps the most significant of such studies - the evaluation of the Bangalore Project is discussed in detail in 2.6 below.

Second phase evaluations emphasise the developmental function of evaluation. This concern is evident in the report on standards for evaluation published by the US Joint Committee on Standards for Educational Evaluation (quoted in Weir and Roberts 1994: 20-24), which called for “evaluations that have four features, *utility*, *feasibility*, *propriety* and *accuracy*, ... in that particular order”. This clearly underlines the shift in focus from accuracy - internal and external validity, and reliability - in the view set out above. Combined with this new role for evaluation, there is a new methodological

paradigm as well. Pawson & Tilley (1997) cites the centrality of craft here, represented in the work of Patton (1982; 1995) as involving “skill; discipline; clarity; creativity; competence; care; etc” (1997:13). It also brings into the process of evaluation, a set of questions regarding use of the findings. *Developmental* evaluation theory is therefore concerned with the reporting of findings, the framing of critical debate on these findings, and ultimately the taking of appropriate action (Rea-Dickins 1994a). This functional, pragmatic approach to evaluation is represented in the ELT field in publications such Rea-Dickins & Germaine (1992), Weir and Roberts (1994), and publications related to evaluation practice such as Mackay (1994a); Mackay *et al* (1995); PRODESS Guidelines (Kiely *et al* 1995).

The third phase is represented in the evaluation theory of Norris (1990) and Pawson and Tilley (1997). Here the concern is not just with accuracy or functionality, but with explanation: the task for the evaluator is to explore and establish links between the specific components or innovations in the programme. These links inform on “patterns of reasoning and volition” (Pawson and Tilley 1997:57) across the project context, and provide the policy-maker, not with a simple answer to cause and effect questions, not with description of the contribution of programme inputs, but rather an account of how the innovations have led to changes in the programme. With this *explanation* remit, evaluation in this phase assumes the purpose of research as well as its methodology. It is the study of programme implementation, potentially different from the programme design or rationale. This approach to evaluation is evident in one particular strand of language programme evaluation: the work of Mitchell (1989; 1990; 1992). She combines a strong theoretical orientation with a grounded qualitative approach, which seeks to represent the process, and account for it theoretically, as a requirement for answering the specific questions of policy-makers.

This three-phase view of theory development in evaluation, summarised in Figure 2.2 below, is important for this study in a number of ways: first, the diversity of perspectives on evaluation process and purpose provides a useful backdrop for

	Main evaluation purpose	Research Strategy	Evaluation quality criteria - how the evaluation can relate to policy development
Phase 1 Evaluations	To determine whether the programme is successful	Experimental study Hypothesis testing	Internal validity External validity
Phase 2 Evaluations	To contribute to development of the programme	Multi-faceted inquiry Goal-free Hypothesis generation	Functionality (Utility Feasibility Propriety Accuracy)
Phase 3 Evaluations	To explain how the programme works	Multi-faceted inquiry Hypothesis generation Hypothesis testing	Explanation Impact on programme theory

Figure 2.2: Summary of three conceptions of evaluation

understanding the perceptions of those involved in the programme evaluation studies in this thesis. Second, phases 2 and 3 conceive of programme implementation as different from programme design, a view which corresponds to recent work on how teaching differs from curriculum or syllabus design (e.g. Woods 1996) and is especially useful in the context of programme evaluation by teachers. Third, the developmental focus of phase 2 evaluation emphasises the central role of practitioners: i.e. those who can effect improvements to the programme. Fourth, the explanation remit of phase 3 evaluation illustrates how evaluation contributes to policy development through understanding of how the programme actually works. Finally, each phase conceives of an audience, increasing as additional functions are included: Phase 1 - managers and policy-makers; Phase 2- managers and policy-makers, and practitioners; and Phase 3 - managers and policy-makers, practitioners, and those with an interest in the social sciences.

2.3 Domains of programme evaluation

In addition to the chronologically-structured influences on ELT evaluation theory and practice, there are influences from three different domains in which evaluations are carried out, and what might be labelled *communities of practice* established.

- a) Evaluation of social programmes in education and other fields where public resources are invested;
- b) Evaluation of foreign and second language education programmes;
- c) Evaluation of English language teaching programmes.

First, more evaluations are carried out in a) than in b) or c) as these incorporate a huge amount of public sector activity and investment. Language programmes are a relatively small group of the public sector social programmes which are routinely evaluated. For example, the Burstall study of French in the primary school curriculum (Burstall *et al* 1974) was, by the early nineties, the only publicly-funded, large-scale evaluation study of foreign languages in the UK primary curriculum. One impact of the low volume of programme evaluation in language education has been a difficulty in the development of a programmatic approach, i.e. a series of linked studies, to the issues which relate to language programme success of otherwise. This in turn has resulted in a reliance on language tests developed in the psychometric tradition of evaluation as indicators of overall programme quality; and on classroom research studies in the psycholinguistic tradition investigating SLA questions.

Second, there is a shared tradition of applied research and evaluation in the work carried out in education a) and b) - which is not so evident in c). This is illustrated by the profile of work of a university school of education, and by the perspectives on evaluation/research studies, as Norris points out “The way we think about evaluation is heavily conditioned by different conceptions of applied research and its relationship with the politics of educational administration and policy” (1990:11). In the specific context of language programme evaluation, the idea of a shared research and

evaluation agenda permeates the discussion in the papers in the special second language programme evaluation edition of the journal *Language, Culture and Curriculum* in 1990 (Harris; Mitchell; Ullman; Swain *et al*; Gardner *et al*). This linkage differs from the more applied, context-bound evaluations which characterise a substantial proportion of evaluations of English Language programmes. These describe a range of programme activities without engaging with a cumulative body of theory. Bowers (1983); (Swales 1989); Mackay (1994b) label such evaluations as *War Stories and Romances* to represent the descriptive aspect of such evaluations. The discussion below on the nature of programme theory in ELT explores why it has been difficult (for evaluators and others) to develop a cumulative body of theory, or indeed draw upon the existing knowledge in social science and educational evaluation.

Third, a good deal of ELT programme evaluation has been carried out in the education systems of countries other than that of the sponsors and evaluators: programmes which are part of aid projects have evaluations commissioned by the Department for International Development (*DfID*, formerly *Overseas Development Administration – ODA*) of the UK Government and *The British Council*, and carried out by UK evaluators. Evaluations often have to be carried out in a manner which looks at the English Language programme, but not at the educational policy context or management culture within which it is implemented. The effect of this distance between programme evaluation and educational policy has been to concentrate on ‘neutral’ theoretical issues, and issues of practice, without engaging with issues of educational policy and values¹. This feature of ELT evaluations has meant that evaluation research has not been able to engage with the issues of attitudes and values which proved so important in foreign and second language programme evaluation work within the UK.

Fourth, in ELT there is a significant tertiary and private sector within which programmes are designed and implemented. These are ‘evaluated’ by market forces, rather than by programme evaluation processes which characterise publicly-funded

education programmes. There is a certain amount of external evaluation through inspection mechanisms, e.g. British Association of State English Language Teaching (BASELT) and British Association of Lecturers in English for Academic Purposes (BALEAP) recognition schemes where the focus tends to be on institutional capacity to deliver programmes rather than on the nature of specific programmes. This is evaluation in the inspection tradition, increasingly with an important developmental purpose (Wilcox 1992; Blue and Grundy 1996). There is little engagement with theory building, as recognized by the Applied Linguistics and Education fields; rather the inspection process may be seen as constituting an alternative body of programme theory, professional rather than intellectual, which often sits ill with the latter. This particular divide is evident in the debates in the professional journal *ELTJ* – Swan (1985) and Widdowson (1985); Prodromou (1996a; 1996b) and McCarthy and Carter (1995; 1996); Senior (1997a; 1997b) and Hatfield (1997).

In conclusion, this range of influences on evaluation practice within the ELT context has provided a diversity of conceptions of evaluation, and of approaches to implementing them, but limited empirical verification of suitability of approach for task. And without a range of studies investigating which features of programmes contribute to learning and how, the programme theory is only partially developed. It tends to be characterised by strong theoretical foundations, particularly from Applied Linguistics, well developed structures for practice - the craft elements of classroom practice - but weak on policy issues: how a programme in a given context can be designed, implemented, understood and improved. The issue of student-student interaction in the language classroom illustrates this: there is a strong case for including such interaction in the curriculum from SLA theory (for example, Long and Porter 1985), from educational studies (for example Barnes 1976); there is a range of task types and published materials which teachers can use to promote interaction, and the training of teachers routinely draws attention to techniques for interactive activities. There are, however, few published studies which explore how such a curricular component is appropriate and successful in different contexts - different

types of learners such as young children; different physical contexts such as large classes; contexts with specific goals such as examination classes; specific cultural or ideological contexts such as classrooms in lockstep educational traditions, etc².

This account of evaluation activity in different educational domains illustrates the problem for the ELT evaluator: there is a wealth of ideas about evaluation, and for carrying out evaluations, but little empirical evidence to signal the issues to focus on. A further complication is the two dominant research traditions which influence both conceptions of programme and investigative activity in the ELT field.

2.4 Evaluation in ELT - Two research traditions

Theoretical input into the design and implementation of English language teaching programmes comes in equal measure from two very different epistemological traditions: linguistics and applied linguistics on the one hand, and curriculum processes within education on the other. These traditions embody a diversity of research methodologies, but two have been particularly influential in shaping the paradigm debate in programme theory in ELT. On the one hand, experimental and quantitative methods from psychology became in the 1970's, the dominant approach in the study of SLA within applied linguistics, most notably by Krashen and colleagues (Krashen 1982; Krashen and Terrell 1983). On the other, the work of Stenhouse and others in general education in Britain in the same period highlighted the role of the teacher in developing the curriculum in the classroom through negotiation (Stenhouse 1975). This was complemented by the work of Barnes who linked this epistemology of the classroom to the role of interaction in learning (Barnes 1976). These two traditions of programme research and development resulted in a common design - communicative language teaching. The key features of this classroom methodology are still very much part of the orthodoxy: learning through communication, involving contexts of language use relevant to the learners in the classroom.

This conjunction of quite different ideas in programme design poses two problems for evaluation. Firstly, what kind of knowledge is the evaluation of the practice trying to inform - a psycholinguistic account of second language acquisition, or an account of pedagogic processes in the classroom. Secondly, which research methodology is most appropriate for the examination of programme practice: a quantitative approach which assumes that an adequate description has been elaborated, or a qualitative approach which holds that the process is not describable in generalisable terms, and the research task is to describe in context in such a way as to account for the contribution of the different features. These problems have defined the paradigm debate in evaluation in English Language Teaching. In many ways, this debate corresponds to the classic divide between experimentalists seeking to determine, in a generalisable way, which factors in a programme cause effects, and naturalists constructing an account where the goal is to establish links between different elements of a particular programme.

2.5 Programme theory in ELT

A feature of social programmes is that they are always based on a programme theory - "a set of understandings about how certain inputs and actions lead to certain results" (Pawson and Tilley 1997:215). The programme theory may be elaborated formally in the traditional domains of sociology, psychology or anthropology, or it may be more akin to a theory of practice as described by Schön (1983): knowledge from experience that there are patterns to human actions and reactions. It is interesting that Schön's work on reflective practice has been a significant influence in the development of teaching and teacher education in ELT over the past decade. His account of professional action derives from his thesis that those implementing programmes are not just rolling out a plan designed by experts - technical rationality - but rather are acting out a tacit theory - knowing in action - or are working with a more aware account of praxis - reflecting in practice (Schön 1983:50-69). This analysis is especially pertinent to programme theory in ELT, where, as discussed above, the domains of theory and practice often seem remote from each other.

A basic problem in understanding evaluation in ELT is the lack of consensus over what comprises a programme. At one level the answer to the question “What is a good language programme?” has two answers - one reflecting theory which describes a defensible construct, and another reflecting practice, which looks to what succeeds in the market place. A quick glance at the papers presented at a major professional conference such as IATEFL illustrates the range of components which informs programme theory in ELT: analysis of syntax, cognitive accounts of learning processes, sociological perspectives on groups and classroom interaction, neurolinguistic programming, the role of drama and the verbal arts, ethnography for the language learner, quality assurance procedures from management studies, etc. This diversity of sources of information and inspiration prevailing in ELT practice testifies to the complexity of language teaching, to the difficulties teachers face when trying to achieve language competence outcomes. The parameters of the theory in the field have traditionally been more limited and perhaps, manageable, but the effect here is reductionist - a narrow defining of the field such that programmes with theoretical integrity can be designed, and evaluated.

The notion of theoretical integrity of foreign language programmes was established with Audiolingualism, in the 1940's and 50's. Skinner (1957) set out the basic theory of learning - behaviourism - and the principles of contrastive analysis of the first language and the target language which could generate a programme of learning for the classroom and the language laboratory, with little or no planning or creative contribution from the teacher. Chomsky (1959) demonstrated that language form and language use were too complex to be learned thus, starting the new search for a language programme theory. In the communicative era, new understandings about language use and language acquisition led the search for the theory, with Wilkins (1976), Widdowson (1978), Canale and Swain (1980) and Breen and Candlin (1980) being key milestones. However, the essential coherence of Audiolingualism became the defining shape of a programme theory. It characterises Krashen's Natural method

(Krashen and Terrell 1983), derived from his empirical work on the Monitor model, and perhaps even more seminally, determined the nature of method elaborated by Richards and Rodgers (1986).

Richards and Rodgers' analysis exemplifies the *technical rationality* approach (Schön 1983). They introduce a core area of basic theory - of language and of learning, an elaborate design field where the real world factors are systematically taken into account, and finally implementation in terms of teaching acts and learning acts (Richards and Rogers 1986). It is an account of programme theory with the needs of the programme evaluator in mind - there is a chapter about comparing methods, a central concern of a programme evaluator in the technical rationality/positivistic paradigm. Richards (1984) develops in some detail the relevance of this analysis to programme evaluation. Yet, as the discussion of the evaluation of the Bangalore Project (see 2.6 below) illustrates, it is the experience of the evaluators that illuminates most clearly the shortcomings of this programme theory.

Theoretical work on communicative language teaching also generated basic programme theories during the 1970's. Some accounts such as Munby (1978) proved limited and short-lived. Others, such as Widdowson (1978), Canale and Swain (1980) and Breen and Candlin (1980) have proved more enduring, and have been the basis for a series of debates about the nature of language programmes throughout the 1980's and 90's. As Legutke and Thomas (1991) point out, however, these programme theories are essentially a set of hypotheses which need to be tested in the real world of language courses and classrooms. The specific contribution of Legutke and Thomas is to set out a description of language teaching, where practice in terms of teaching and learning acts is related to the key theoretical constructs of the communicative curriculum. This programme theory differs from that of Richards and Rogers in two key respects. It starts with the holistic nature of educational development and the place of language learning within that. Second, it accords a strong role to the learner, the autonomous, aware learner whose *process competence* -

capacity to participate in and benefit from curricular tasks - is a key element in the design and implementation of the programme. All participants in the programme are evaluators - making judgements about the value of the learning tasks, and taking developmental action as part of the process³. These features of the programme theory point to the requirements for programme research and evaluation: it is a task of description and explanation, informed by all possible perspectives on the classroom experience.

In addition to Legutke and Thomas' account, one might list a range of published coursebooks, professional groups' inspection schemes, institutional programmes⁴, or individual teachers' approaches as programme theories within ELT. The experience of evaluation in the field illustrates the diversity of conceptions of the components of ELT programmes, as illustrated in Figure 2.3 below.

The first point about these conceptual frameworks is that they are designed for evaluation purposes. Rea-Dickins, in a review article, is focussing on the components of programmes that have been the focus of evaluations. In the case of the Brazilian ESP project the framework represents a merging of what teachers and others might consider as components of a programme's success, and what can be determined through data. Mackay is elaborating an evaluation approach to on-going quality management and assurance. In the case of Weir and Roberts, the conceptual framework relates closely to calendar time, so that findings can be factored in to a cycle of programme development. A common component to all is *teachers/ teaching/ classroom management*, though from the other elements in each framework, it is clear that there is no shared view of what teaching actually involves: how does teaching relate to materials? how does it relate to learning outcomes? does it include tasks, or are tasks within the category of materials? The discussion on classroom observation in evaluation in ELT below looks more closely at how teaching is conceived of as a component of programme evaluation.

Evaluation Framework	Components
<i>Rea-Dickins (1994)</i> <i>[broad overview]</i>	<ol style="list-style-type: none"> 1. Teachers 2. Materials 3. Projects, programmes and courses
<i>Celani (1988) / Alderson & Scott (1992)</i> <i>[programme within a specific language centre]</i>	<ol style="list-style-type: none"> 1. Attitudes 2. Approach/methodology 3. Classroom Management 4. Materials 5. Learning outcomes 6. Impact of outsiders 7. Teacher training 8. Publications 9. Research by teachers 10. Administration 11. Resource Centre
<i>Mackay (1994a)</i> <i>[programme within a specific language centre]</i>	<ol style="list-style-type: none"> 1. Quality of each course 2. Quality of teaching 3. Student performance
<i>Weir and Roberts (1994)</i> <i>[programme within a specific language centre]</i>	<ol style="list-style-type: none"> 1. Objectives 2. Student language achievement 3. Materials 4. Teaching 5. 'Customer satisfaction'

Figure 2.3: Four evaluation accounts of ELT programme theory

The concept of *programme* is central to this study, and is investigated empirically as an integral part of the task of researching programme evaluation. Chapter 5 explores in detail what constitutes the programme evaluated. The next section looks more closely at the practice of language programme evaluation in recent decades. The starting point for this discussion is the evaluation of the Bangalore CTP project. The four evaluation issues profiled by this evaluation are then traced through a range of evaluations in the 1980s and 1990s. These are: the role of experimental design; the measurement of outcomes; the description of the component parts of the programme; and the involvement of the programme participants in the evaluation.

One reason for the importance of the Bangalore evaluation to the development of evaluation in the English Language Programme field is the extent to which the evaluation issues have been documented by one of the evaluators - the discussion here draws on Beretta (1986a; 1986b; 1986c; 1987; 1989a; 1989b; 1990a; 1992a; 1992b), the related Applied Linguistics / ELT literature - Greenwood (1985); Brumfit (1984a; 1984b); Prabhu (1987), and the evaluation literature where the Bangalore evaluation is considered in detail - Rea-Dickins and Germaine (1992), Weir and Roberts (1994), and Lynch (1996).

2.6 The Bangalore evaluation and ELT evaluation research

The Bangalore project developed from a second language pedagogy theory elaborated by Prabhu from theoretical work on language use, and a psycholinguistic account of language learning. The key feature of this theory is that a second language would be most successfully learnt through exchanging meanings, rather than through a focus on grammatical forms. The teacher's role is to engage learners in the classroom in a series of tasks which involve the expression in the target language of meanings already clear through another medium - the classic example being the railway timetable, a database from which a range of verifiable propositions can be articulated. Classroom activities involve cognitive manipulations of the database, with a distinct drilling structure, and interaction is teacher-learner only (lesson transcript exemplifying this methodology is set out in Prabhu 1987).

This programme and its evaluation are important in ELT for a number of reasons:

- It was constructed round the idea of a complete teaching method (See Section 2.5 above), and thus was seen by curriculum theorists (Brumfit 1984a) and practitioners Greenwood (1985) as relevant to their task;

- It involves the application in the classroom of a developed, mainstream body of theory (Prabhu 1987), set in the mainstream psycholinguistic tradition of the study of SLA;
- It placed the notion of task at the centre of communicative language teaching, and contributed to a sustained discussion on the nature of tasks for language learning;
- It constitutes an example of implementation of task-based learning in ELT, in an era when TBL is the subject of widespread inquiry and research (Candlin and Murphy 1987; Nunan 1989b; Skehan 1996)
- It was implemented at a point where there was a widespread sense of change in English language teaching, based on a single set of methodological principles (learning through communication) supported by developments in the theory of communicative functions (pragmatics, speech act theory) in British Applied Linguistics (Wilkins 1976; Widdowson 1978; Munby 1978) on the one hand, and by developments in the theory of naturalistic second language learning in American Applied Linguistics (Krashen 1982; Krashen and Terrell 1983) ;
- It was supported by a prestigious sponsor in English language teaching - The British Council, and through this support the project was visited by key people in the development of theory and practice, which ensured a significant impact on ELT conferences and publications.

An evaluation of the project was commissioned:

To assess through appropriate tests, whether there is any demonstrable difference in terms of attainment in English between classes of children who have been taught on the CTP and their peers who have received normal instruction in the respective schools.

Prabhu, quoted in Beretta and Davies 1985

The evaluation was carried out, using an experimental, comparative design, in the final months of the project. The discussion below examines four aspects of the evaluation approach, as evident from the responses to the evaluation, and from the evaluation literature published in the following decades.

a) The role of experimental design

The experimental approach to evaluation of social programmes has long been seen as the most objective, and ultimately permitting the most generalisable conclusions. In ontological terms, the approach takes the view that social processes mirror processes in the natural world - the agricultural-botany paradigm in Parlett and Hamilton's seminal critique, which "utilises a hypothetico-deductive methodology derived from the experimental and mental-testing traditions in psychology" (Parlett and Hamilton 1972:2). Once the variables in the programme to be evaluated have been described, and a method for quantifying their effects has been elaborated, a comparative study can proceed. This issue is important because it represents an orthodoxy in social programme evaluation: it has a credibility which is important for those who commission evaluations. The key implication for evaluations in ELT is the view that one classroom can be considered the same as another, except for the specified variables, which distinguish experimental from control classrooms. In the Bangalore evaluation, the specified variable was *method*: experimental classrooms used the task-based, procedural syllabus, while the control classrooms used conventional grammar-based teaching methods and materials.

The Bangalore evaluation focussed on outcomes, with an assumption that the outcomes relate to two different teaching methods: the evaluators sought a design "which would be experimentally valid, [..with..] tests which would be equally fair to both teaching methods" (Beretta and Davies 1985: 122). In their rationale for this approach to the evaluation, Beretta and Davies acknowledge that previous method comparison studies "failed to yield conclusive results". They quote Stern:

Chapter 2

The inconclusiveness of these studies does not mean that the research is a waste of time. The studies gradually revealed that the 'methods' are not clearly defined entities that can be juxtaposed and compared. It would be a waste of time if that lesson had not been learnt.

Stern 1983: 71

Beretta later (1986a;1992b) acknowledges the lesson about methods:

results from such an inquiry [a testing comparison] would not be very helpful. They would neither indicate that a teaching method worked, nor even that the teaching approach existed.

Beretta 1992b:264

Two major surveys of programme evaluation in ELT - (Beretta 1992a; Rea-Dickins 1994) consign the experimental approach to the archives. Alderson (1992) in discussing the issue of experimental design concludes that while it may have a role in research projects, it is not appropriate for evaluations in ELT⁵:

Although their aims were clearly evaluative in nature, their time span, the nature of the questions being asked and the resources invested in the projects make them appear less relevant to the sorts of evaluations that are typically commissioned.

Alderson 1992: 283

Pawson and Tilley's (1997) critique of experimental design has a clear ontological basis: the outcomes of social interventions are never wholly predictable, and any evaluation of a treatment which compares predicted outcomes is likely to provide inconsistent results. The human response to a given situation is mediated by a number of individual and contextual factors, and if these are not explored in the context of the evaluation, the measures of predicted outcomes versus others are not very useful. It amounts to a reliance on *successionist* rather than *generative* logic (1997:32), an expectation that people behave in response to a social situation as predictably as litmus paper responds to an acid or alkaline solution.

Harris defends the positivistic, experimental approach, particularly for evaluations which are "more concerned with causal questions and theoretical issues and favour an

experimental/quasi-experimental approach with the focus primarily on product data” (Harris 1990:83)⁶. Harris sees multi-faceted approaches - essentially qualitative process data - as relevant to the evaluation of a programme when the “programme represents something of a black box to the concerned agencies and/or to the evaluator” (1990:86). It is interesting to note that Harris in the same paper proposes that the experimental product approach is useful in precisely this situation: “when in the context of a local programme evaluation, it is difficult for political reasons to study certain kinds of performance effects (e.g. pupil attitudes to learning a particular language)” (Harris 1990:86). This view illustrates the issues of black boxing in evaluation, and the nature of language programmes. Black boxing is the reliance on assumed descriptions of processes where precise accounts are not available. These descriptions can derive from one aspect of the theory underlying a programme, and exclude others. An example of this in the case of the Harris evaluation is the issue of teacher action: the focus of the research is the behaviour of language learners, so what the teacher does, and why, remains marginal. It is black-boxed - an assumed description of the teacher’s reasoning and behaviour is set out and considered to apply in all classrooms. Harris, discussing the evaluation illustrates this reasoning:

When the initial national surveys on spoken Irish were being carried out, and the main evaluation report was being prepared, the audio-visual *Nuachursai* had been almost universally used and accepted since the mid-1960’s, so that there was a fairly widespread understanding, both inside and outside schools, about how Irish was taught and what went on in a typical Irish class. The fact that the *Nuachursai* emanated from the government department of Education, whose primary school inspectors were all subject matter experts in Irish, added to the feeling that the teaching of Irish proceeded along well-recognized lines. Thus, while it would have been useful for the purposes of our earlier evaluation to have had various kinds of process data available, such data probably would have been considerably less important than they might be in other situations, at least as far as decision making and policy formation relating to the programme itself was concerned.

Harris 1990:87

The central assumption of Harris here reflects on the one hand, the discussion of language programme theories (Section 2.5 above), and on the other, studies such as the Lawrence evaluation (Section 2.6c below). This assumption - that teaching is according to guidelines issued by the Ministry of Education - is a dubious one, and unlikely to inform judgements about the programme in a useful way.

A strong argument for experimental designs is set out in FitzGibbon (1996), based on evaluations carried out using public examination results in a range of British secondary schools. However, three features of the evaluation approach she advocates represent key differences from the practice in language programme evaluation. Firstly, the 'experiment' is ongoing. She writes about collating data over time using sophisticated information technology systems in secondary school contexts such that there is a rich backdrop against which to interpret findings. Secondly, "fair" and "really fair" comparisons require engaging with value added elements - residual data - within each context, a methodology for taking patterns of value added into account when designing comparative evaluations. Thirdly, the function of such evaluations is to support internal development within "the unit of responsibility". Thus, in contexts where extensive data are collected over time and measures are taken conscientiously to improve the quality of these data, an experimental approach is useful. Such a comparative approach to evaluation for internal development might be considered relevant to the context of evaluation studied in this thesis: in relation to higher education programmes, there is an easily accessible database of module and programme assessment grades. Taken together with student profiles and teacher evaluations, these data could be used for development of the programme by programme teams and departments. This issue is explored further in Chapter 8 below.

b) The measurement of outcomes

The purpose of the Bangalore evaluation was to determine whether there was any demonstrable difference in the attainment of two groups of learners, experiencing two

different kinds of language programmes - CTP and normal instruction. To achieve this the evaluators:

decided to construct a battery of tests intended to measure *achievement* separately for experimental and control groups (by a structure test and a CTP task-based test), and *proficiency* by three neutral measures: contextualised grammar, dictation, and listening/reading comprehension.

Beretta and Davies 1985:123

The use of tests in programme evaluation in this way is based on the “the proof of the pudding is in the eating” logic - a language programme has the purpose of generating language learning, and it is appropriate to judge the worth of that programme by determining the level of proficiency at the end. The final judgement of the evaluation in the published report, after all the limitations are considered is that the test results:

provide tentative support for the CTP claim that grammar construction can take place through a focus on meaning alone.

Beretta and Davies 1985:126

Since there is no account of what “a focus on meaning alone” actually meant in terms of the programme as implemented, the tentative support must be considered very tentative indeed. While my argument here raises no queries with the validity of the tests designed, there is a the more general problem of the validity of the link between test results and different programme activities.

Lynch (1996) points to another drawback of a reliance on the measurement of learner outcomes in programme evaluation: one where the strategy is to use pre- and post-tests with a single group rather than post-tests with experimental and control groups as in the Bangalore evaluation.

In those contexts where [...] no comparison group exists, quantitative analysis is limited to non-parametric tests of the significance of pre-test - post-test differences. Without a comparison group, of course, the observed change from pre-test to post-test is difficult to interpret - it may have occurred in spite of

the programme, or an even greater change may have resulted from an alternative programme.

Lynch 1996: 105

Tests have maintained their role in ELT evaluation, but with changes in emphasis. Rea-Dickins and Germaine (1992) describe tests as evaluation instruments which primarily give teachers and learners feedback on the effectiveness of learning activities, rather than determine the success of a whole programme. Thus, measurement of learning outcomes is appropriate for smaller units - of time or topic - within a programme, because the information can be used to fine-tune and develop classroom activities. Such tests, designed and implemented by the teacher are likely to be strong on internal validity, in that they relate closely to the learning focus of the programme. Aggregation and standardisation of these small sets of test data are an essential part of educational processes. At tertiary level in Britain, the outcomes of these processes are quickly converted to national league tables of degree classifications, etc. This phenomenon may impact negatively on the development potential of such data (FitzGibbon 1996), and thus compromise the consequential validity (Messick 1994) of the testing process as a whole.

The evaluation process described in chapters 3 and 5 of this thesis identifies a clear divide between evaluation and assessment. The latter is seen as a mechanism to allow students to progress or otherwise, rather than a means of knowing on the success or otherwise of the programme the students have experienced. This view represents in part the emphasis on process at the expense of outcomes. It is a view which, taken to extremes, is problematic, since it excludes both an important perspective and data set from the evaluation process. Chapters 5 and 6 explore this construction of evaluation and assessment in the context of one particular programme.

c) The need to describe the component parts of the programme

Beretta's comment after the evaluation that it was unclear "that the teaching approach existed" (1992b:264) underlines the need for the evaluation process to engage with

the task of description. Without this description of how the construct of the project - Prabhu's method - translated into classroom action by teachers, stakeholders could not interpret the findings: sponsors did not know in what ways their investment worked. Practitioners remained uninformed on issues of "teacher and learner performance" (Greenwood 1985) in the classroom, and theorists did not get the empirical underpinning of the principles of task based learning.

Attention to how programme constructs are operationalised is a feature of key evaluations post-Bangalore. For example, the Lawrence evaluation of the methodology of Zambian teachers of English (reported in Rea-Dickins and Germaine (1992), and Lawrence (1995)) developed a conceptual framework from the main pillars of communicative language teaching, which permitted a mapping of the actual curriculum. This evaluation explored the extent to which there was a match between the needs of the learners (in terms of proficiency in English in an English as medium of instruction context) and "the appropriacy of a syllabus in relation to the context in which it is used". From an examination of the curriculum literature in general education, a series of four continua were used as a conceptual framework for locating ESL teaching between traditional/structural and communicative poles. The four continua are:

Perception of language in syllabus:	synthetic - analytic
Function of language in the classroom:	use - usage
Nature of language practice activities:	cognitive - mechanical
Nature of language teaching strategy:	Inductive - deductive

Adapted from Lawrence (1995)

This framework facilitated the description of the curriculum using data from teacher questionnaires, teacher interviews, and systematic observation in classrooms.

In terms of evaluation theory development in the ELT field, the Lawrence study makes a significant contribution: it provides a comprehensive framework for investigating ‘communicativeness’ in the language curriculum, and the implications of this for effectiveness and efficiency. Recent work on the contribution of teachers to programme design (Graves 1996; Woods 1996; Bailey and Nunan 1996) which borrows from work in teacher thinking research in general education (Clandinin 1986; Connolly and Clandinin 1988; Calderhead 1988) shows how what actually happens in classrooms can differ from the planned curriculum. Thus it is necessary to examine, not only what happens, but the intention of the teacher, the rationale that guides a teacher’s response to a given situation. This suggests a more open approach, along naturalistic or ethnographic lines, to ensure that all data is included, not just that accommodated by the pre-ordinate categories which, for example, Lawrence distilled from her construct analysis.

The approach for description of the classroom process developed by Lawrence is part of an extensive technology for systematic classroom observation developed during the 1970’s and 1980’s. Wallace (1991:66-83) describes these as either in the Bellack tradition - systems based on discourse analysis of classroom language - or in the Flanders tradition - systems based on real-time documenting of observable behaviours in the classroom. The approaches developed for different evaluations by Mitchell and colleagues (Mitchell *et al* 1981; Mitchell 1992) build on both these traditions, to develop a pedagogical account of what happens in the classroom as well as an analysis of classroom language (see Section 2.3 above).

The multi-faceted methodology developed by Mitchell for the evaluations of French and Gaelic language programmes within schools in Scotland seek to take into account the complexities inherent in the classroom process, and in the wider social context. Mitchell (1989) discusses a range of evaluation studies in the Stirling tradition - evaluations of language programmes by Mitchell, and her colleagues Parkinson and Johnstone, which might be characterised as the “systematic and qualitative” study of

language classrooms and programmes. The aims of these evaluations have a clear descriptive and explanatory element, for example, 'To document current instruction practice ...'; 'To document the attempts of committed teachers to implement a communicative approach'; 'To explore the potential of using the target language'; 'To develop an operationalised model of communicative competence ...' (Mitchell 1989:196). The approach to description has a basic research goal, with the judgmental function located as an optional application. Mitchell (1992), for example, states in relation to the evaluation of the Bilingual Education Project (BEP) in Scotland:

In undertaking the project, however, the evaluators themselves expected not so much to solve one specific policy question, as to provide substantial and detailed accounts of the workings of bilingual education, and of the contextual parameters which appeared to constrain or promote it, which could form a more general background to policy making by others.

Mitchell (1992:124)

This approach illustrates shift in evaluation function from Phase 2 - a focus on recommendation to improve policy and practice - to Phase 3 - explanations (see Figure 2.2 above). The theorising implicit in the explanatory function, however, finds the descriptions generated by systematic classroom observation wanting. Mitchell (1989) sees the framework for the systematic study of language classrooms which she and the Stirling school developed as providing only a partial account of what is going on in classrooms and within programmes. The approach suffers from a number of limitations common to systematic analysis instruments. It is dependent on five key assumptions:

that a common set of categories can be applied to produce a valid picture of a series of unique events (lessons and incidents within them);

that an observer can "read" the intentions of the teacher regarding the kind of language experience being provided for his/her pupils, from inspection of his/her overt behaviour;

that pupils in classrooms normally interpret ongoing activity in the way that the teacher intends, and are usually focussing on those aspects of language which are sanctioned in the official lesson plan;

that that naturally occurring speech events can legitimately be categorised as having either a given feature, or not having it - multiple codings or interpretations are not permitted.

... that its only unit of analysis is a relatively “coarse” lesson segment.

(Mitchell 1989:199)

Mitchell’s conclusion is that such systematic observation instruments “are likely to remain too blunt a tool for these investigations to prove ultimately fruitful” (op. cit. 1989:200). Her account of the BEP (Mitchell 1992) illustrates the array of political, cultural and personal factors and sectoral interests involved in a language education programme, which the evaluator must i) negotiate in order to carry out the evaluation, and ii) describe, if the evaluation is to inform the policy decisions which have to be made in relation to such programmes. This view illustrates the potential of qualitative study of programmes for theorists, policy-makers and teachers alike, and informs the methodology elaborated for the study of programmes in Chapters 3 and 4 of this thesis.

One well-documented evaluation in ELT which sought to describe the contribution of stakeholders such as teachers from their own perspective is the Brazilian ESP project (Alderson & Scott 1992). The principal strategy here for describing the component parts of the programme - a strong role for all participants in all stages of the evaluation - is discussed in the next section. The process of evaluation described in Chapter 3 below also relates to this problem of describing the programme, in its attempt to account for what the programme is, as well as describe the evaluation activities and findings.

d) The need to involve the programme participants in the evaluation

Participants are usually involved in the evaluation of programmes as informants, i.e. data providers, not as stakeholders. In the Bangalore evaluation this meant pupils providing examples of language in tests, and teachers providing lesson and questionnaire data. From there it is the task of the evaluators to organise, analyse, interpret and report these data. The weaknesses implicit in a reliance on learning outcomes, and the need to describe the component parts of the programme have together led to the view that greater attention should be paid to the perspectives and experiences of the participants. Such involvement has a theoretical and pragmatic basis. For evaluation to be democratic and empowering (Macdonald 1976), and actually used for programme development (Weiss 1986; Hopkins 1989; Murphy 1996; Murphy and Rea-Dickins 1998) it has to represent the voices of different stakeholders. There may be a problem when the voices are discordant and conflicting - the problem of using what seem like a series of very different constructions of the same programme (Kushner 1996). However, it is clear from the empirical studies in this thesis - see Chapters 3 and 5 - that any programme is a collection of different experiences. Pawson and Tilley (1997) propose a solution here: their realistic evaluation documents all experiences, and “tries to follow how the programme enters the subjects’ reasoning” (1997:189).

Related to this is the pragmatic rationale: if changes to practice are required, these will only be possible if the participants such as teachers are persuaded that changes are worthwhile. Engagement with the data and the arguments - critical debate - is perhaps the most successful means of achieving this persuasion. Change comes from a realisation that change is necessary and appropriate. Alderson and Scott (1992) provide a clear account of what such participation might involve:

sharing the decisional, planning roles as well as the donkey work amongst all involved, ... doing a fair share of the planning and a fair share of the donkey work, all participants gaining benefit from the work involved.

Alderson and Scott 1992:38

This is difficult to achieve: there are issues of power and hierarchy in programme contexts which make such democratic involvement in evaluation difficult (Beretta 1992b; Murphy and Rea-Dickins 1999). In addition, managers and evaluators have the time and skills to develop the evaluation whereas practitioners such as teachers have a range of other routine duties. Beretta (1992c), in a postscript to the Alderson and Scott account raises other issues: the rigidity of a participative approach, the ownership of the data, the definition of participants or stakeholders, and the old problems of bias, whether considered a problem with the outsiders or insiders. Involvement of programme participants in the various phases of an evaluation has traditionally been seen to involve subjectivity and advocacy, and the threats these pose for reliability - practitioners are naturally going to see as most appropriate what they have done, and defend positions they have taken (Beretta 1989). The exclusion of participants on the other hand, can be seen as a threat to validity - particularly where, as in the programmes studied in Chapters 3 and 5 below, teaching is seen as the construct of the teacher, shaped as much by personal, sectoral and strategic interests as by theoretical design.

Traditionally the division of labour in ELT programme evaluations has been a perspective on the practice from the practitioners such as teachers, and a perspective from the theory from the external evaluator - a Jet In Jet Out Expert from a department of Applied Linguistics or Education. The importance of critical debate and recommendations for development in the future (Rea-Dickins 1994a; Weir and Roberts 1994) have meant that programme evaluations should be carried out by programme specialists rather than evaluation specialists. Pawson and Tilley (1997) suggest an alternative division of labour: the external evaluator provides expertise in evaluation skills, working with the insiders to make explicit the programme theory, the patterns of reasoning and volition which underpin what is actually being done. The logic of this approach suggests that recommendations for change or development come from the programme rather than generated by canons of theory from distant academies.

2.7 Programme evaluation by teachers

The discussion above of the development of four aspects of evaluation theory - the role of experimental design; the measurement of outcomes; the need to describe component parts of the programme; and the need to involve participants in the evaluation - over recent decades illustrates two problems. First, despite clear analyses of shortcomings and innovative designs to meet these, the achievement of the purposes of evaluation is still difficult. Second, there is a lack of clarity about the nature of the language programme being evaluated: a focus often on what should be done in the classroom rather than what is done; on curriculum aims rather than on teacher intentions and learners' experiences. The context of evaluation which is the focus of this thesis - programme evaluation by teachers - is one which provides an opportunity to explore these areas.

Evaluation by teachers has been driven by two trends in the development of social programmes in recent decades. First, the greater accountability required of programmes using public funds has led to monitoring tasks which cannot be met by traditional inspection routines. This trend can be seen in the education field in the evolution of practices for monitoring the quality of higher education - the Quality Assurance Agency (QAA) relies less on actual evaluation events, and more on evidence of internal evaluation procedures. Increasingly, teams of teachers in the university sector are required to evaluate their programme using the four grades in six programme areas of the current HEFCE framework, or to submit a Self Assessment Document (SAD) for internal periodic review. Blue and Grundy (1996) demonstrates how the professional body accrediting EAP programmes equally looks for evidence of ongoing internal processes of evaluation and quality assurance. This trend for accountability through internal monitoring is also evident in other fields - Cooper (1995) illustrates how various charity programmes are working to achieve similar goals through internal participant-led evaluations.

Secondly, there is a new understanding of the applied nature of social programmes. The constructivists (e.g. Lincoln and Guba) and the realists (e.g. Pawson and Tilley) are in the ascendant, demonstrating that programmes are not the blueprints handed down to practitioners for implementation, but are constructed on the ground, informed and shaped by the set of contextual factors prevailing, negotiated by those involved. While evaluation can be seen as a tool for teachers to develop this kind of curriculum, in the empowering, Stenhouseian manner, it can also be a control mechanism (Parton 1994 – see Chapter 1 above). This characterisation of evaluation as a potential ‘spy’ in the classroom, disempowering rather than empowering, a tool of management rather than one for teachers, is one echoed by teachers involved in this study, and revisited in Chapters 3, 5 and 6 below.

Programme evaluation by teachers tends to be based on methodologies received from the evaluation contexts discussed above. The literature (e.g. Rea-Dickins and Germaine 1992; Weir and Roberts 1994), sets out the ways in which evaluation can be done to discharge the two established functions of evaluation: accountability and development. Increasingly these concerns are seen as components of quality, as different aspects of the processes of quality enhancement and quality assurance. For example, the developmental imperative has also shaped the perspective on evaluation of the policy-makers and sponsors - validating bodies such as BASELT and BALEAP, and project sponsors such as The British Council increasingly look for evidence of evaluative activity within programmes to make them more effective (Blue and Grundy 1996; Rossner and Strange 1995; PRODESS Publications 1993-1997). This policy level support for developmental evaluation borrows extensively from a tradition of teacher research for curriculum and professional development within education stemming from the work of Stenhouse and colleagues on the Humanities project, Hopkins (1985; 1989), Gitlin and Smythe (1989). It is supported by recent work in teaching and syllabus design in communicative language teaching, which describes the process nature of English language programmes, locating quality firmly

in the management of that process (Legutke and Thomas 1991; Graves 1996; Bailey and Nunan (Eds.) 1996; Woods 1996).

However, while the purposes of evaluation by teachers and methodological guidelines are clear in themselves, they represent a tension in four ways. First, carrying out an evaluation is characterised as research, posing basic skill and resource problems for teachers. Hammersley (1993) argues that teaching and research are essentially different activities, requiring different skills and orientations to the classroom. Second, evaluation methodology revolves round specific questions, whereas teachers cannot easily move away from a necessarily holistic view of the classroom and the programme. The programme evaluation by the teacher described in Weir and Roberts (1994) reflects this holism, and the evaluation policy described in the next chapter essentially places teachers in a responsive role in the evaluation process, pursuing issues raised by the classroom rather than responding to their own pre-conceived questions. Third, the developmental imperative requires a speedy, often real-time approach to data processing and analysis, and to critical debate and action which is counter to established research and evaluation methodologies. Fourth, the interface with orthodox research remains undeveloped. The process-product research role for small scale evaluations by teachers in real teaching contexts as described by Long (1984) and raised within the action research framework by Crookes (1993) does not seem to have been successful, and the divides between the theory and practice epistemologies in English Language Teaching seem as wide today as at the time of the Widdowson/Swan debate in the mid-eighties.

2.8 Other traditions of teacher inquiry

The factors explored above which have contributed to programme evaluation by teachers have also promoted other forms of teacher inquiry in their classrooms: action research; reflective practice and classroom ethnography, and the more overtly management process of teacher appraisal. These investigative procedures have very different origins, but share aspects of method and purpose with evaluation. They are

examined here for two reasons: to ensure a comprehensive overview of teacher inquiry in ELT; and to identify aspects of programmes which these traditions focus on, but which have received limited attention under the evaluation label.

a) *Action Research*

The origins of Action Research lie in research in management studies and social psychology, for application in industrial contexts. Lewin's approach to study of social and organisational processes emphasised the local:

Lawfulness in social as in physical science means an "if so" relation, a relation between hypothetical conditions and hypothetical effects. These laws do not tell what conditions exist locally, at a given place at a given time. In other words these laws don't do the job of diagnosis which has to be done locally. Neither do laws prescribe a strategy for change.

Lewin 1946, quoted in Adelman 1993:11

Crookes (1993), defines two principal forms of action research for the study of English language programmes: the first kind is research done by teachers, to test theory-driven hypotheses elaborated by conventional researchers. Long (1984) and Cohen and Mannion (1995) reflect this view, the former labelling it *process evaluation*. Action research thus only differs from conventional research in the fact that the researcher is a participant in the process.

Action research of the second kind is:

"..... teachers analysing the way their own practices and understandings are shaped by broader ideological conditions [and] ... linking reflection to action, offers teachers a way of becoming aware of how those aspects of the social order which frustrate change may be overcome".

Carr & Kemmis 1986:179-80

In this model action research requires teachers to take a critical stance to existing educational policy, hierarchical relationships and prevailing epistemology handed down

from the academy: "Foreshadowing and engendering a different form of social organisation" (Crookes 1993:133). At one extreme, this represents teachers as oppositional or anarchic (Gibson 1985; Elliot 1991). Interpreted more moderately, it requires teachers to examine the wider context of their pedagogy, and be aware of the implications of their innovations for prevailing hidden policies and unwritten rules (Kemmis and McTaggart 1981; McNiff 1988; Somekh 1989; Zeichner 1993).

A central feature of this model is the dominance of data over theory. As Usher and Bryant (1989:74) put it "Practice is *underdetermined* by theory" in classrooms, so there is a need to look afresh at practice, to examine whether new paradigms might not yield a better explanation of pedagogic and schooling processes. This feature is resonant of the essential empiricism in evaluation theory discussed in 2.6 above, and of the enduring problem of conceptualising programmes as design constructs - an issue examined in detail in Chapter 5 below.

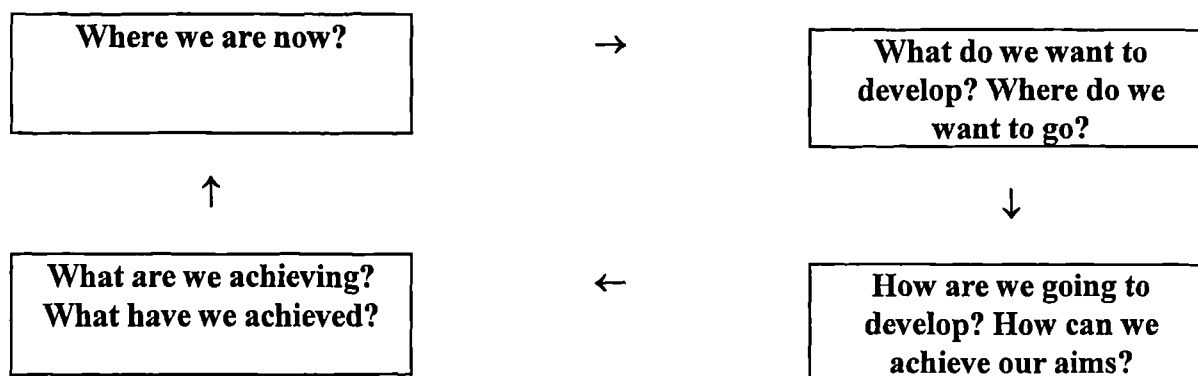
The similarities between action research and classroom evaluation have been recognised and built on to guide teacher inquiry. Rea-Dickins and Germaine (1992) characterise action research as:

.... evaluation undertaken by teachers and other ELT professional for themselves in their own particular settings.

Rea-Dickins and Germaine 1992:70

They set out the cyclical frameworks for inquiry initially developed by Kemmis and McTaggart (1981) and McNiff (1988). In a later paper one of the authors (Rea-Dickins 1994b) represents evaluation as ongoing and cyclical in the manner of action research:

Chapter 2



Rea-Dickins 1994b:4

A more fundamental and complex parallel between action research and evaluation is set out by Norris (1990). His discussion of the social and moral bases of evaluation brings together the 'critical self-reflection' of Habermas, and the concerns for democracy and morality of Macdonald and House, and suggests an evaluation which genuinely questions the appropriateness of policies and their ideological foundations, and seeks to empower all participants so that their interests are given equal voice. This approach to teacher inquiry has been applied in mainstream education in Britain, for example, the work of the Classroom Action Research Network (CARN), supported by a community of academics and evaluators at the University of East Anglia (Elliot 1991; 1993; Kushner 1995; Norris 1998). However, in the ELT field this strong characterisation of the local remains an aspiration rather than an actuality. As discussed above, programme theory has developed around clearly articulated orthodoxies, and evaluation has been concerned with maintaining these rather than critically examining them.

b) Reflective practice

Reflective practice has its origins in the study of the underlying, tacit knowledge of practitioners in different professional domains (Schön 1983). This work originally was concerned with how basic knowledge from the subject area combined with experiential knowledge and creative thinking together enabled professionals to solve new problems in a principled way. This work has come into ELT as a model for the training of teachers, where a broad concept of reflection is seen as a means of relating

theory (from Applied Linguistics) to the practice of teaching (Wallace 1991; Richards and Lockhart 1994). In general education, reflective practice celebrates the experiential knowledge of established teachers, tracing its philosophical roots back to the work of Dewey (Pollard and Tann 1987).

Whereas the theoretical characterisations of reflective practice by Dewey and Schön describe tacit, implicit learning and programme improvement, the applications for language programmes emphasise explicit learning through classroom investigations similar to action research or evaluations (Wallace 1991; Richards and Lockhart 1994; Wallace 1998). The focus is on teachers in training, and the reflection is guided, in the case of Wallace by data from the classroom, and in Richards and Lockhart, by theoretical frameworks from Applied Linguistics.

Reflective practice in a general sense remains a powerful and positive label for sound professional practice in the ELT field. It is not part of Schön's or Dewey's argument that reflective practice is what should be done, that institutions and professional bodies should establish a policy for reflective practice, or seek strategies to use communally the results of the reflection. Rather, reflective practice celebrates both the inevitability and the private nature of teachers' learning through practice. The account of teacher development in the classroom described in Chapter 5 below reflects these features, while the manner in which this learning is embedded in an evaluation processes illustrates the need to approach teacher inquiry from different theoretical perspectives.

c) *Classroom ethnography*

Classroom ethnography is the application of a tool for developing understanding in anthropology to the study of classrooms and other areas of educational activity.

Ethnography developed to identify features of the social order where there is no explicit rule system, and to explain behaviour of various participants in such contexts. All behaviours, including language, are parts of a social semiotic which can be

explicated to reveal fundamental aspects of our human and social nature. The study of classrooms within this paradigm can thus facilitate understanding of routines, and changes in routines - especially useful in the context of innovation (Holliday 1994; Roberts 1998). While ethnographies generally inform on the nature of social interaction, major studies in the field of second language education claim a direct contribution to the understanding of intercultural communication, and the development of teaching and learning methodologies (Van Lier 1988; Roberts *et al* 1992; Holliday 1994).

The use of ethnography in evaluation has been to counter the orthodox measuring of performance within programmes according to pre-ordained measures, and to provide a basis for explaining patterns of events which at the outset are poorly understood, for example, Kiely (1996) and Dubin and Wong (1989). This approach to evaluation represents a theory from data epistemology, for the purpose of redressing power imbalances in established hierarchies, and allowing evaluation to critically examine issues of policy and ideology as well as observable categories of performance. Holliday (1994; 1995) uses ethnography to explore patterns of programme success and failure which would remain inaccessible to other evaluation approaches. Ethnography is better established as a research method than an approach to evaluation, although there is a body of practice in American educational evaluation (Fetterman 1988; Guba and Lincoln 1989; Stake 1995) where ethnographic approaches are used to generate grounded programme theories (see Chapter 4 below).

d) Teacher evaluation in the appraisal tradition

One function of evaluation in education has been to monitor the work of teachers. Although such activity can have a positive developmental role - Wilcox (1992) describes how inspectors and advisors can collaborate with teachers to improve teaching and learning - much of the evaluation within this tradition has been controlling and limiting, concerned with narrow views of teachers' effectiveness. Gitlin and Smythe, surveying both the history of teacher evaluation, and the assumptions which underpin

contemporary schemes for teacher appraisal, set out six critiques of this monitoring form of evaluation and its key assumptions:

- i) it involves a separation of knowing from doing, formulation from realisation, theory from practice;
- ii) it takes a pseudo-scientific approach, taking behaviours into account, while not considering intentions;
- iii) it constructs evaluation as a technique, elevating instrumentation over other aspects of the process;
- iv) it denies accumulated professional knowledge, claiming teachers only know what the prevailing policy states, and can only act within that framework;
- v) it is a monologue, where the teacher is advised by an expert on what is needed for improvement;
- vi) it constructs teachers as minor players in the educational process, enhancing the contribution of the institution, and of experts outside the classroom to the curriculum at the expense of teachers.

Gitlin and Smythe 1988:27-37

Gitlin and Smythe's analysis develops from the study of teacher evaluation processes in general education, but is useful here for four reasons. First, it provides a frame for understanding the role of policy-making bodies which set out criteria for teacher qualification and programme recognition in ELT. Second, it establishes a critical theory for examining the role of institutional policy in shaping the curriculum - in the programmes studied in this thesis, there is a complex tension between teacher autonomy and institutional policy, both in the teaching of the programme, and the evaluation. Third, it provides a view of evaluation where the teacher is not invisible, as in, for example, the Bangalore and Lawrence evaluations described above. Fourth, it sets out theoretical underpinning for the development of horizontal evaluation, a developmental and empowering form of teacher evaluation, which realises the key principles of an evaluation process which is democratic (Macdonald 1976) and critical in the Habermas

tradition. Gitlin and Smythe set out four key requirements for horizontal evaluation, which has as its goal to foster change by enabling teachers to base change on a critical accounting of the ethical and political consequences of the curriculum. First, the evaluation is collaborative and dialogic, where different interpretations of what happens in classrooms are explored. Second, each participant accepts that his or her view is partial, or even wrong, and can be made more complete or accurate through dialogue. Third, a critical friend, enters the dialogue on the same basis as other participants, and introduces new factors which can develop explanation and understanding. Fourth, the relationship between intention and practice is analysed at the outset, and this analysis is revisited and clarified during the evaluation dialogue.

The approach of Gitlin and Smythe shares philosophical foundations with the theoretical perspectives of evaluation, action research and reflective practice. In addition, their view of teacher evaluation shares many features with the institutional scheme for teacher appraisal in the institution where the programmes studied for this thesis are taught: there is a peer appraisal scheme in the institution; the peer is identified by the teacher and takes on the role of 'critical friend'; professional performance and development over the preceding period is evaluated without recourse to set benchmarks; and development goals for the forthcoming period are negotiated. This appraisal scheme is separate from the programme evaluation policy, and programme evaluation data are not used in the appraisal process. The reasons for and consequences of this divide are explored through empirical data in Chapter 5 and 6 below, and an 'integrated' solution is set out in chapter 8.

2.9 Teacher inquiry - summary

The traditions of teacher inquiry for classroom development discussed in this chapter - evaluation, action research, reflective practice, classroom ethnography and teacher appraisal - differ in respect of their philosophical origins, epistemological underpinnings and communities of practice in education. The differences from one context to another, tend to relate to the following parameters:

collaborative	—	individual
teacher-initiated	—	manager-initiated
emancipation	—	effectiveness
grounded	—	programmatic
one-off	—	ongoing

Figure 2.4 Features of different teacher inquiry traditions

For example, evaluation practice as explored in this thesis tends to reflect the right-hand column, although the approach is grounded in the student experience, rather than based on received criteria (see Chapters 3 and 6 below). Action research of the second kind (Crookes 1993 – see 2.8a above) tends to correspond to the left-hand column, although it might be on-going, a series of linked studies as represented for example by McNiff (1988).

These differences often obscure the fact that all forms of teacher inquiry have a shared purpose – to improve teaching and learning by taking stock of what happens in the curriculum – and thus a shared role in the process of quality management and curriculum development. Chapter 8 of this study explores how this notion of *shared role* might be developed, such that the type of inquiry described in the intervening chapters might have greater impact.

2.10 Summary

The evaluation literature as surveyed in this chapter is suffused with potential, but also cognisant of the degree to which this is often unfulfilled - The Bangalore Evaluation; the Lawrence evaluation in Zambia, The BEP evaluation; the Brazil ESP project. In all these cases the guiding theory raised expectations of improved policies and practices. In all, the reality is one of unfulfilled potential. Development is one purpose of evaluation. The other is accountability. In relation to publicly-funded programmes the concept of

accountability has its roots in a moral axiom: that in a democracy the citizens have a right to know how public funds are used.

In recent decades the axiom has become linked to the concept of quality, and evaluation in different forms – formal research-type studies, feedback from stakeholders, inspection, teacher inquiry and appraisal – is the process by which quality is monitored and managed. It is this multi-functional form of evaluation, represented diagrammatically in Figure 2.5 below which places the classroom teacher in a central role in the process, and which is the subject of this thesis. The strategy for exploring this subject is an empirical one: to describe the unfolding of a programme, and the role of evaluation within this.

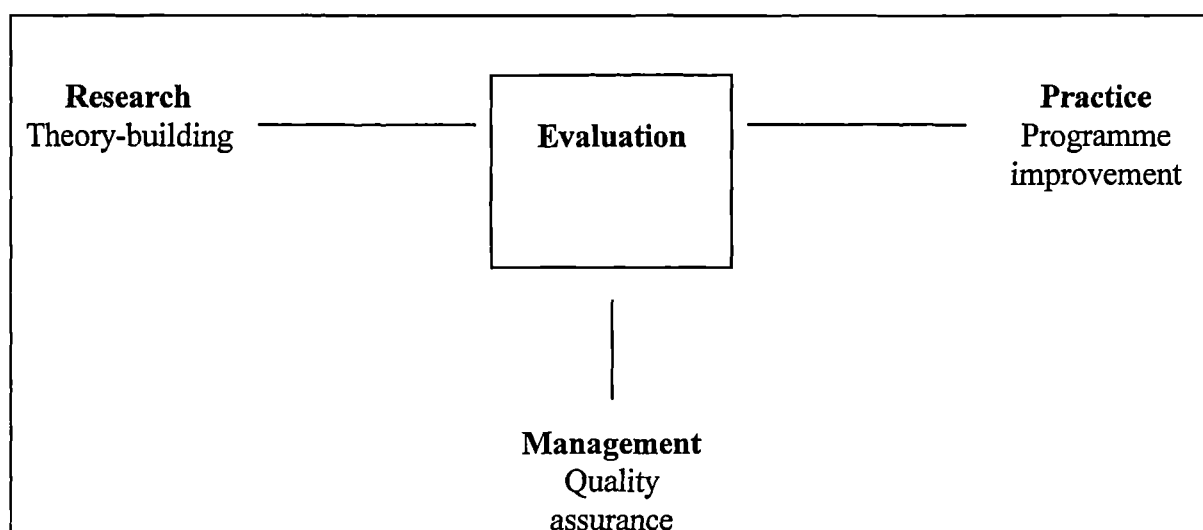


Figure 2.5: Multi-functional evaluation

One aspect of the research task in this thesis is descriptive: using a grounded approach to determine what is involved in the process labelled ‘evaluation’, and carried out by teachers within their own programmes. This description task generates three research questions:

RQ1. How do teachers manage the different agendas - their own personal and professional interests, students concerns, institutional requirements - involved in carrying out evaluations?

RQ2. To what extent is involving students in evaluation a separate process from teaching them?

RQ3. How do teachers use the time available to carry out the evaluation and deal with other pedagogic demands?

A more critical perspective deriving from the issues set out in this chapter and the preceding one, considers the validity and impact of the evaluation process. This involves the links between the programme as construct and the evaluation; the feasibility of the evaluation policy in terms of teacher and student roles, and the worth of the evaluation: the extent to which it enhances learning, teacher development and quality, as conceived by the wider community. This focus is framed in four research questions:

RQ4. To what extent are the evaluation findings valid and reliable?

RQ5. To what extent do evaluations focus on problems and explain them such that improvements are brought about?

RQ6. To what extent is the experience of others - relevant literature - used in investigating problems and devising solutions?

RQ7. To what extent do evaluations influence the pedagogy in terms of classroom activities, materials, assessment procedures, etc.?

Chapter 2

The next chapter sets out a pilot study of programme evaluation, exploring and developing these questions.

Chapter 3

Pilot Study - The evaluation of an EAP programme

3.1 Introduction

The discussion in the preceding chapter illustrates the multiplicity of discourses underpinning and informing evaluation in the context of English language programmes. These discourses include a range of purposes for evaluation which span quality management and improvement of practice as well as the more conventional research function. This chapter studies the evaluation of an English for Academic Purposes programme in a British university. First, the institutional and policy context is set out, followed by a brief analysis of the aims of the programme. Then, the evaluation processes within the programme are examined. Finally, the research questions outlined in the preceding chapters are revisited. The discussion examines how they can be refined in order to better facilitate empirical exploration of this aspect of the curriculum in the main study for this thesis.

3.2 Institutional context

The institutional context can be described at two levels: a British “new” university¹, and within that, a department with an established record and reputation in developing English language and related teacher education programmes. The university is an expanding institution (characteristic of the sector in the mid-90’s), with an evaluation policy which seeks to improve the quality of modules and programmes on the one hand, and discharge formal accountabilities to stakeholders on the others. The policy (see Appendix 6) outlines ‘general principles’ of evaluation, and describes the flow of information from teaching teams to Directorate. This process has a development goal - each team and unit is required to take action on problems and disseminate ‘best practice’ - and an accountability function - the information flow underpins the institutions quality assurance systems (see 6.4.a below for a detailed analysis of this

policy). This policy, however, is separate from policies and procedures on assessment, research, programme design and review procedures, and staff appraisal - an issue revisited in Chapter 8 of this thesis.

Two quality audits by professional bodies² in recent years commend the concern for quality in the department, and comment on the role of evaluation in this. BASELT reports from 1992 and 1995 trace the development of evaluation policy and practice in the department:

A course committee is responsible for the quality of each course. Evaluation from students is formalised for the full-time courses, but apparently not well-organised for the part-time courses. The Department is conscious of the need to formalise this as well and is looking for ways of obtaining feedback.

Overall the assessors recommend that a more formal evaluation of the course should take place in order that both the admissions policy and the design of the course itself are meeting the needs of the students. This evaluation should take place on a formal in-course basis and also through more formal follow-up procedures from former students studying on courses.

BASELT Report 1992:5

[The institution] has quickly put into place an excellently clear and unfussy quality management framework, curriculum evaluation is very well embedded in module and course processes. Formative evaluation using structured group feedback, is carried out mid-module and the results are summarised for students and the programme leader and useful agreed changes are implemented.

BASELT Report 1995 para. 4.5

For the HEFCE Teaching Quality Assessment (TQA) in 1996, the time when the programme evaluation described in this chapter was carried out, the role of and approach to evaluation in the curriculum was set out (see Figure 3.1). This approach illustrates the focus on the interaction between the module as approved by the institution, the teacher or teaching team, and the students, both during the programme and at the end.

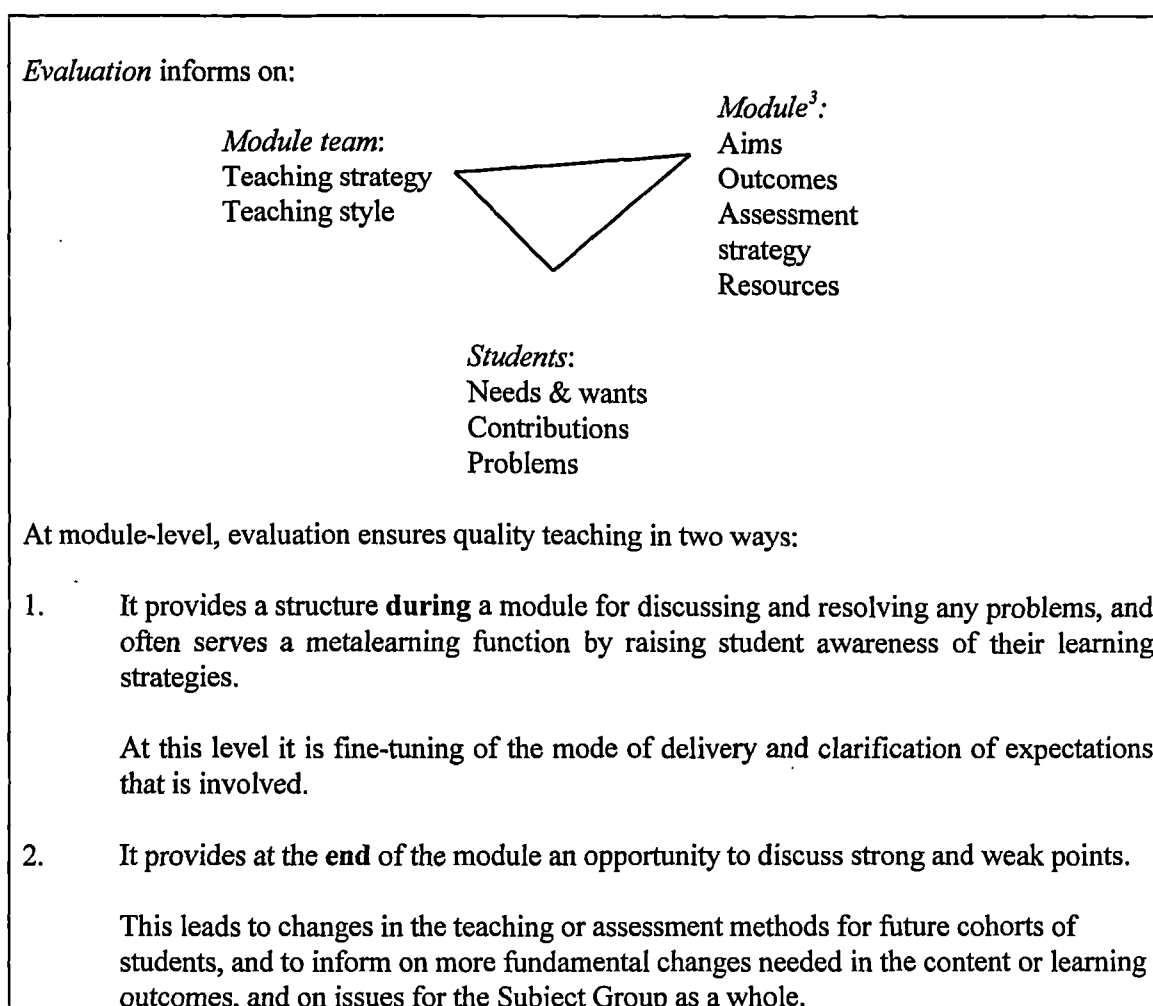


Figure 3.1 Departmental approach to evaluation of programmes

The evaluation policy was reported on within the Quality Assurance (awarded four out of four) section of the TQA report:

Students have a voice in the evaluation of modules and programmes offered by the Department. There is a range of evaluation procedures across the subject groups from open discussion to written evaluations in which the students in the early part and at the end of the semester make their opinions known. Issues identified in module evaluations [.....] are acted on.

HEFCE Quality Assessment Report 1996 para. 51

The ‘open discussion’ procedure was the Nominal Group Technique (henceforth NGT). This evaluation procedure, a structured focus group (Kreuger 1994), was initially developed for improving health care programmes (van de Ven 1972; Delbecq 1975). It is used in both mainstream educational and ELT evaluation (Wilcox 1992; Weir and Roberts 1994), and in higher education to develop a democratic approach to measuring student satisfaction (Green *et al* 1994). Appendix 6 includes the NGT procedure recommended for modules within the department.

NGT was adopted as a departmental policy for mid-module evaluation for a number of reasons: the link to language learning; time efficiency; opportunity for learner training; going beyond ‘It depends’; avoiding evaluation fatigue; facilitating teacher development (Kiely 1999a). These views developed in part from a lack of enthusiasm in the department for an institution-wide questionnaire on student satisfaction (see Appendix 6 IV.E), and an awareness of the wider experience with that approach (Marsh 1987; Richardson 1994). The NGT was thus seen as acceptable to teachers, and also, as evidenced by the quality audits discussed above, capable of meeting quality management procedures. It represents the principal ‘formal’ evaluation in the two studies for this thesis (3.5; 5.3 below), and is revisited in the discussion of evaluation policy in 8.2 below.

This body of evidence illustrates both the coherence and credibility of the evaluation policy of the department, based on the programme improvement discourses outlined in Chapter 2 above. It also shows that the policy was implemented, that evaluations were carried out, and seen as having an impact, also documented in publications such as Clark *et al* (1996), O’Regan (1997) and Kiely (1999a; 1999b). However, the place of evaluation in the curriculum is not uncontested. While teachers implemented the evaluation policy, and used the data and reports for the centralised quality management purposes, three reservations were frequently voiced in relation to evaluation. First, the requirement that teachers provide regular reports of student feedback to line managers was seen as indicative of a view that teaching is a technical

enterprise that needs to be carefully monitored and controlled. Second, the reporting was a bureaucratic task which took time away from the teaching activities which actually benefited students. Third, the reports could constitute a performance review of teachers, for use by managers, but outside the institutional appraisal framework.

The case study reported in this chapter, seeks to document how in an actual programme, these positive and negative dimensions of evaluation are realised: to what extent does the evaluation process support classroom practice?; and to what extent do the negative elements constitute barriers to improved teaching and learning?

3.3 The programme

The programme is a thirty-six hour English for Academic Purposes (EAP) course with two 90-minute classes a week over twelve weeks. Two teachers, Millie and Anna⁴, teach a class back-to-back each week, sometimes overlapping or team-teaching. The students - fourteen students are registered, and attendance fluctuates between five and twelve, with an average attendance of nine - are non-native speakers of English from undergraduate programmes, such as Accountancy Studies, and Information Management. The broad aims of the programme (see 5.1 below) focus on helping students make the best use of existing knowledge and skills in English in their studies, and encouraging them to explore effective ways of meeting the linguistic demands of their chosen programme of study. The assessment procedures involve assessing student performance on oral and written presentations in English, the latter in examination conditions as well as through coursework. Students are given marks out of a hundred, which go to the examination board, and thus impact on their academic progress in general.

The teachers, Anna and Millie, are part of a team of four teachers of EAP programmes. To provide a backdrop to the study of the evaluation of this programme I asked all four teachers to return detailed questionnaires, focussing on their professional biography in relation to evaluation, and their attitudes to, and practices

within the current policy framework. The questionnaire is attached as Appendix 2. The findings show that the teachers learn about evaluation through a process of experiential induction, rather than as part of their teacher education experience. All teachers expressed the view that the evaluation policy they were working with was valuable, especially the NGT procedure. The value, noted by both Anna and Millie in the questionnaire, and also expressed in interview by Millie, and in the classroom by Anna, was the opportunity it provided for language skills practice. The focus on this particular feature of NGT raises important questions. Is evaluation only of value to students insofar as it provides a context for language skills development? Is it the case that the understanding generated about teaching and learning is of value to the teacher only? What does this view say about the process of negotiation at the heart of their pedagogy? These questions emerging from the initial questionnaire study of the EAP team, together with the more general critique at departmental level contrast with the positive view of evaluation represented by the quality audits. This contrast is the starting point for problematising this form of evaluation, and underpins the empirical study of it.

3.4 Research methodology

Evaluation can be considered a policy-culture conflict, where policy represents the centralised approach to quality management, and culture, the autonomous way in which teachers manage their classrooms and programmes⁵. A research strategy therefore, should document the evaluation process in its wider programme context in such a way that key phenomena can be identified, and understood in a way which facilitates a more focussed approach for the second (and main) case study in this thesis. Thus, the goal is to “catch the complexity of a single case” (Stake 1995:xi), and relate this complexity to aspects of the wider problem. The research design was informed by what Guba and Lincoln label naturalistic inquiry:

Naturalistic inquiry is always carried out, logically enough, in a natural setting, since context is so heavily implicated in meaning. Such a contextual inquiry *demands* a human instrument, one fully adaptive to the indeterminate

situation that will be encountered. The human instrument builds upon his or her *tacit* knowledge as much as if not more than upon propositional knowledge, and uses methods that are appropriate to humanly implemented inquiry: interviews, observations, document analysis, unobtrusive clues and the like. Once in the field, the inquiry takes the form of four successive elements: purposive sampling, inductive analysis of the data obtained from the sample, development of grounded theory based on the inductive analysis, and projection of the next steps in a constantly emergent design. The iterations are repeated as often as necessary until redundancy is achieved, the theory is stabilised, and the emergent design fulfilled to the extent possible in view of the time and resource constraints. Throughout the inquiry, but especially near the end, the data and interpretations are continuously checked with respondents who have acted as sources, as well as with counterpart individuals; differences of opinion are negotiated until the outcomes are agreed upon or minority opinions are well understood and reflected.

Lincoln and Guba 1985:187-188 (italics in original)

The approach to observation was informed by the tools of ethnography:

The ethnographer participates, overtly or covertly, in people's daily lives for an extended period of time, watching what happens, listening to what is said, asking questions, in fact collecting whatever data are available to throw light on the issues with which he or she is concerned.

Hammersley and Atkinson 1983

My task was to 'lurk and soak', in order to develop a 'thick description' of the process of evaluation, such that the phenomena of enlightenment, adaptation and betterment could be documented, and the links between them illuminated. I was concerned to work with an *emic* rather than an *etic*, i.e. to document the programme and evaluation as it is conceived by the participants, especially the teachers, rather than according to some external construct.

Collecting data involved field notes, scribbled furiously in the classroom to both describe actions and capture snatches of dialogue, and head notes to augment the field notes from memory recorded immediately after each lesson. These were typed up and

given to the teacher concerned to check within a day or two. This observation data was complemented by four other procedures:

- questionnaire studies of the four teachers involved in this kind of programme;
- detailed interviews with Millie, the teacher mainly focussed on;
- ongoing informal discussion of the programme and observation accounts with Millie and Anna;
- examination of all documentation from guidelines to students to students' work.

Research design and strategy issues are revisited at the end of this chapter, and examined more fully in the next chapter.

3.5 The evaluation of the programme

The approach to the evaluation taken by the teachers was largely within the institutional framework discussed in 3.2 above. The focus was on student self-assessment of needs, then on student perceptions, collected through the NGT procedure, and at the end of the programme by means of an open-ended questionnaire.

a) *Stage 1 - Baseline evaluation - establishing the pedagogy*

The first four sessions of the programme each included a baseline evaluation activity, labelled, inter alia, needs analysis⁶ and self evaluation. In the first session the students did a writing task for diagnostic purposes, and received two sets of guidelines for the programme. The *Module Guide* (5 pages) provided procedural information - contacting teachers, calendar, assessment procedures - and different foci for lessons: *language areas*; *study skills*; and *course content topics*. - studying abroad, travel & tourism, television and young people, etc. The *Study Guide* (10 pages) was produced for a wider group of students, and contained a 'consumers' charter' section, detailing rights and responsibilities for students; a section on academic writing conventions;

and a section on independent learning, introducing the self-access facilities available, and strategies for using the local community as a learning resource. The discussion was relaxed, encouraging students to participate by contributing their particular query or concern.

Students talk about their needs in relation to the module guide: Abe, the student sitting near me, says that when writing, he just thinks from one word to the next, never of the ideas taking shape.

Anna: These are problems to do with your situation in British HE - we will help, but all the language problems cannot be eliminated here.

Another asks about getting help with assignments for other modules. Another asks about interpreting assessment tasks, and other lecturers' feedback.

Teachers [Anna and Millie] indicate intention to help with these areas, insofar as time and fairness allow.

Classroom notes CS1:1⁷

In the second session, students use a worksheet to identify the learning needs vis-a-vis sixteen skills. They rate the importance for them of each skill, rate their competence in that skill, and then outline an action plan to improve their performance in five of these. The third session revisits this worksheet with a plenary discussion of skills and action plans. The fourth includes a reading of the resources section of the study guide, with Millie outlining which resources might be used to resolve particular skills issues.

The data illustrate a pattern in these classroom discussions: a student identifies a real problem; the teacher identifies self-access materials which might help, states that the input in class will deal with the problem, and limits expectations due to the constraints of time within the programme. A second, related feature of these sessions was the teachers' attempts to persuade learners to engage in independent, resource-based learning. Millie, for example, is aware that this persuasion task is a challenging one, and feels that while it is a major element of the 'formal' pedagogy, it is not an aspect of learning which she can integrate into her teaching.

Later Millie expressed dissatisfaction with students' work outside the classroom - a dead loss: they do not do it, or if they do, they think it is mechanical

Chapter 3

exercises; they do not think about it in the way they should. *She feels that it is not such a good idea to give them work in advance, as they just do not do it, and then the activity planned cannot go ahead. The piece of work later in the session was influenced in the same way - Millie felt that they should have given this more time before the session as there is not enough time to do everything in class.*

Post-lesson discussion with Millie CS1:4

In relation to the triangular structure posited for programmes (see Figure 3.1) the module at this stage can be seen as a construct of the institution and the module team - Anna and Millie. The students, in contrast, seem like consumers rather than negotiators.

b) Stage 2 - Formative evaluation - taking stock and moving on

Anna referred to evaluation in the first session when going through the calendar:

Evaluation comes up in the setting out of the calendar. Anna explains that it is necessary - no reason - and that it will be done through a writing and speaking activity, so that ...

Anna: ... the time spent on it will be useful for what the students are learning. *Seems an uncommitted, apologetic attitude to evaluation.*

Classroom notes - CS1:1

Millie commented (post-lesson discussion CS1:1) that there was a case for including more information about evaluation in the guidelines given, but the prevailing view was that: “we can tackle evaluation best just before we do it” - i.e. when carrying out the NGT procedure in Week 5.

This evaluation, using the NGT (see Appendix 6), involved small group and plenary discussion and was led by Anna, with Millie arriving during the plenary discussion. Anna stated two reasons for the activity:

Anna: ... a university requirement, to monitor quality, and to give students an opportunity to bring about adjustments ...

Classroom notes CS1:5.

Students were asked to note aspects of the programme under three headings:

1. What has worked well for you;
2. What has not worked well for you;
3. Comments

A strong rationale was made for the language learning aspect of the procedure. And from my position in the classroom - behind a group of five students during the small group discussion phase - the activity functioned well as a language learning task: the students were engaged in getting from their own notes to an agreed list of points to feed into the plenary discussion - the report back. The question of adjustments to the programme is less clear cut: the group discussion (see Figure 3.2) shows among other things, the level of generality at which students perceived their needs.

There are four students in the group near me: Carrie, Abe, Adam, and Noah (new student). Carrie starts off by explaining to Noah what they have been doing. She starts explaining what stage 2 is, and asks him:

- Carrie: What would you like more of?
 Noah: Grammar, essay-writing ...
 Carrie: We cannot have too much. How many weeks are left? Only 5?
 Abe: Yes
 Carrie: So we have to concentrate on the most important
 Noah: Writing essays
 Adam: Did you mention grammar?
 Carrie: *(looking at Abe)* And you?
 Abe: Grammar - using tenses. And vocabulary - using vocabulary
 Carrie: Choosing vocabulary? Selecting vocabulary - vocabulary selection
 Abe: Yes
 Carrie: *(looking at Adam)* And you?
 Adam: Punctuation and ...
 Abe: ... prepositions
 Carrie: I agree grammar as well. Essay writing, writing summaries, which is like essay writing. Vocabulary..... *(sound hesitant)*
 Abe: If you don't agree, leave it out
 Carrie: No, no and punctuation
 Abe: Punctuation and some verbs take prepositions, what are they? Like look at, take for, look for,
 Carrie: Ah, phrasal verbs
 Noah: Spelling

Abe:	Teacher can't teach you spelling
Noah:	Well (<i>inaudible, seems to be giving examples</i>)
Abe:	That's quite basic isn't it? Like ei, e before i
Carrie:	Other comments?
Adam:	Frequency of speaking
Abe:	Note-taking
Adam:	Reduce time from three hours to two hours
Abe:	No, no (<i>Carrie nodding agreement</i>)
Adam:	(says something and laughs)
Carrie:	Perhaps break is too long
Abe:	Plus your recommendation
Carrie:	Me?
	(<i>Abe says something I don't hear, which has the effect of closing the discussion.</i>)

Figure 3.2 Group discussion stage of NGT - CS1:5

It seems as if the process of managing the discussion, especially the task of getting agreed conclusions to the discussion which Carrie assumes, is the focus, rather than the analysis of how the programme is meeting their learning needs. The discussion of vocabulary illustrates this pressure for agreement: instead of describing the experience of problems, or probing what might be involved here, the students broaden and generalise the category. They label the problem 'vocabulary selection' but it remains unclear what is involved, or how the programme might be adjusted to meet these particular needs. In the plenary discussion (see Figure 3.3), Anna seems to have little option with the vocabulary issue other than to continue the search for a global label (see 5.5 below for an analysis of vocabulary in the context of CS2).

There are three possible explanations for the apparent limited detail of the students' feedback and lack of insight. It might be that the structure of the evaluation is inappropriate; that the students are unable to further the analysis, or they do not wish to. Beyond the observation that the group discussion does not follow the three-point structure set out by Anna, the classroom data do not inform on the first explanation - a more complete transcript of the group discussion is required to suggest explanations for the development of the discourse. There is evidence to support the inability hypothesis: individual students frequently misunderstood feedback given by the teachers, and tasks

tutors expected to go well proved too difficult. There is also some evidence that students may hold back on their learning difficulties, possibly so as not to represent themselves in a negative light in the group. Abe, for example, seems more interested in commenting on the problems raised by other students, than in probing his own. He describes his vocabulary problem in the classroom as:

Abe: ... using our own words, not using the words in the text.

Classroom notes CS1:5 (see below)

Talking to me (during and after the session) he describes the problem as plagiarism - the feedback he has had from other module tutors. In discussions with Anna, however, in the NGT plenary and in a tutorial the following week, he does not refer to this. The issue of students using the evaluation activity to manage their self-representation within the programme is examined in the context of CS2 in 5.3 below, and in a detailed student case study in Chapter 7 below.

The plenary feedback stage, starts off by examining what went well, but after a brief listing of points on the overhead projector, it became a discussion on preferences for future sessions. Figure 3.3 illustrates how one segment of the plenary discussion becomes a shared exploration of how the learning needs vis-à-vis vocabulary might be met, similar in many ways to the needs analysis discussions earlier in the programme (see 3.4a above). The process is thus oriented towards action for the future almost to the exclusion of evaluation of what has already taken place (see 5.3a for exploration of this hypothesis in the context of CS2).

NGT Plenary discussion: *What did not work well [Heading on OHP]*

Key

A = the student group - one of three

3 = the point proposed, discussed and listed on the OHP - one of five

A - 3

Gina: More skills in writing

Anna: How do you mean?

Gina: Structure with the ideas, we have the ideas, we want to know how to present the ideas

Anna: We need to be more precise, to understand what we can do in the time constraint we have

Abe: Using our own words, not using the words in the text

Anna: Not reading from the text; you want to be able to talk, not with the exact words you have written?

Abe: Yes, but not using the words in the text. Say a summary of a text of 800 words; writing a summary of 500 words in your own words, not with the words in the text

Anna: We do that every week

Abe: We need more vocabulary; more building vocabulary

Gina: That's what I am saying

Anna: How should we do it? (*Silence: all seem to hear this as a question for the teacher to answer, not them, though the teacher may intend it as a genuine question*). Vocabulary directly from the text? We could give you list of words not in context. We could give you lists to prepare

Carrie: We need to remember a lot of different words

Anna: Well, you could keep a note book where you noted all new words that you think you need to learn. We can advise you individually in tutorials, like we will do next week

Abe: If we look at any book - say for the Cambridge exams - we have to choose vocabulary for sentences. Sometimes prepositions for verbs. That's what I am saying

Anna: I'll put down VOCABULARY BUILDING. Some work we can do in class, but you have to use the book list references we gave you

Figure 3.3 Plenary discussion stage of NGT - CS1:5

The data in Figure 3.3 illustrate the search for shared understanding in the classroom. The discussion has two main features: the centrality, even dominance of the teacher and the unease of the teacher with this role. Anna's contributions - requiring greater precision, relating points to existing practice, asking for suggestions, recommending independent learning - shape the discourse, but have the effect of avoiding the issue rather than pinning it down. The students seem to expect Anna to provide information on what they should do to solve their problems, realising their role as consumers of instruction (see 3.4a above), rather than stakeholders co-constructing the programme. The question posed at the end of the plenary discussion illustrates this:

Carrie: What are your recommendations?

Anna: About what?

Carrie: About us. You see our work, what do you think we should do?

- Anna: Our expectations are in the module and study guide.
Millie: What do you mean by recommendations?
Carrie: What would you suggest from seeing our work? (*Carrie then turns to Abe and Adam, and asks*) Is it that?
Anna: Different things for different people. You see it in your written work.
Millie: We will talk about that next week and in the tutorial. Or it may be to do with your self assessment sheet.
Anna: If there is something from your programme, bring that in - work back, something like that. We are working in the dark here - we really do not know what your programme is asking you to do.

Classroom notes CS1:5

Carrie is the spokesperson for this group and is setting out here what Abe said in the group discussion phase (see Figure 3.2 above). In the classroom it was clear that the question struck a chord with all the students - as if it was the question they all wanted the answer to. Anna and Millie's response illustrates the central problem arising from the lack of specificity in the preceding discussion: the insufficiency of the analysis emerging from the NGT procedure as the basis for an action plan, and the need to continue investigation and evaluation on an ongoing basis. The teachers' agenda seems to be to get students to self-evaluate, and work out a personal learning plan - a continuation of the process of persuasion evident in the early sessions. The students' agenda is more complex: each student has his or her own individual preferred outcome of discussions on learning. This agenda is particularly evident in the contributions of Abe, to the different stages of the NGT discussion. Chapters 5 and 7 of this study examine closely the issues of student engagement with different programme activities, especially the evaluation procedure, which have emerged in this pilot study.

The list of action points arising from the evaluation (see Appendix 6 Point 6) was not available the following week. I spoke to Millie about it:

- Millie: Not much has come out of it. Some interesting things, but nothing controversial, they just want more of everything. We've talked about it, but we haven't typed it up yet.

Post-lesson discussion with Millie CS1:6

The report produced some weeks later (passed to me, but not given to the students) included all the items noted on the OHP, with brief discussion of what was positive: “no major points of dissatisfaction” and “the exercise itself seemed to be a successful way of practising some of the oral and written skills which the students had come to the module to acquire”; and what was negative:

Less satisfactory was the fact that most of the student requests appeared to be for practice of activities which were already being offered every week in class. Secondly, there was sometimes no linkage between what they said they wanted to do and any strategy on their part to put in the required effort. This was particularly noticeable in relation to the practice on paragraph and essay writing. These activities were set up and started every week in class and students asked to complete them as out of class activities. Relatively few of them handed in completed tasks on a weekly basis to receive helpful/critical oral and written comments.

Formative evaluation report CS1:8

It is interesting to note the emphasis on writing skills here: this was the weak element in the assessed coursework, with the majority of students getting a fail or very low pass mark. The classroom notes do not reflect this particular issue: neither students nor teachers refer explicitly to the weekly writing tasks.

Three aspects of the NGT procedure are evident from this particular evaluation event. First, the future orientation of the discussion seems to foreclose engagement with what has actually been happening in the programme. Second, the evaluation discourse seems overly focussed on labelling problems, to the detriment of analysis and problem-solving. Finally, the procedure seems to give the students conflicting roles: on the one hand they are passive feedback providers, and on the other, they are responsible for analysing problems and developing appropriate action plans. The data here do not document the student perspective on this programme and the evaluation, thus limiting the scope for exploring more deeply the impact of these conflicting roles. These issues are taken up again in CS2 - see Chapters 5 and 7 below.

c) *Stage 3 - Summative evaluation - preparing for next time*

The summative evaluation was carried out in the last session of the programme. Students were asked to fill in a questionnaire (written by Millie) anonymously which asked for three positive aspects of the programme, three negative, three suggestions for change, and a series of Yes/No questions on use of the resources available. The report on this evaluation, prepared by Millie, listed all the comments made by students in each category. The positive comments relate for the most part to classroom processes, including the teachers' contributions, rather than learning outcomes: only three out of twenty-one comments related to progress made, and two of these were modified:

Techniques about writing paragraphs were very useful. I try to put into practice but still it is very difficult for me to write an essay in the proper style. It takes a huge amount of time to put my ideas in writing.

To analyse different kinds of reading was very useful, but I still need more practice.

Summative evaluation report CS1:12

The negative comments and the suggestions for improvements focussed for the most part on "more of..." and issues of procedure - timetabling, length of break, etc. Responses to the resources section show that one third of students made use of the resources available on a self-access basis. Millie in an interview after the writing of this report, found the comments of the students "unsurprising , bland and uninteresting", in that they concentrated on issues of timetabling and procedure rather than:

...whether the content really suited them; whether it matched up with their needs; problems they were having with their other modules; the issue of what they did outside class; how much help they need to go doing things outside class; how clear it was what they were supposed to do outside class; why they didn't do much outside class in this programme.

Interview with Millie CS1:12

Despite the lack of information on these areas, Millie is positive about the contribution of evaluation to the development of the programme:

- Interviewer: And from evaluation of this module will action be taken in terms of shaping the programme for the next time? Will you do it differently?
- Millie: I don't think I'll do it differently. I think over the course of this programme we have developed some good strategies and some quite good materials, that will continue to develop. So it is not a change of direction, it is perhaps going further in the same direction. Things we tried out this time [are] as a result of evaluation, and our own evaluation of the last programme.
- [...]
- I: You mentioned that one of the things that made this programme different from in the past was the study guide, the materials - study guide and module guide - that you gave the students at the outset. How did this affect how you taught it?
- M: Yes, by setting down and formalizing the idea that taught time was largely to support private study time, looking at the ways in which that was done meant that we put more emphasis on that. I spent more time in developing that. also, we put down how we felt they should be analyzing their own problems, and deciding for themselves on priorities, and signposting them more clearly towards the resources that would address these difficulties. I would say that from the final evaluation, that wasn't perhaps very successful - it never has been - you continually get students saying they don't do a lot of this work outside class because "I'm too busy with reading for other modules, I don't have time".....
- I: The key activity is the self-assessment - using the resources, including the resources in the classroom, such as the teacher - depends [.....] on their developing the idea of what they need. You actually had a very structured and systematic approach to needs analysis, didn't you?
- M: Yes, but it would have needed following up every week. We could have spent the whole 36 hours just on that - a self-access course really - where they come and report back on what they have done outside of class.
- I: And it would have been an individual dialogue in that case. You could not talk to the class about it, because you would just generalize beyond each problem then.
- M: But you might be able to. You might see recurrent themes. We can bring some strands together here, and discuss them as a group. There are obvious advantages in that - getting students to compare notes on how they support learning outside class; they can even learn things from one another.

Millie is aware of the challenges posed by her pedagogic approach, both in terms of the students' capacity to engage with it, and the limited time available. There is a sense at the end of the programme of its being part of something larger, of a contribution to the development of her pedagogic approach, which will benefit other programmes and students in ways too imprecise to fit into the action points section of a programme report. The programme over, the students cease to be the centre; they take their place as experiences which tell a story, as frames for devising new solutions for similar programmes in the future.

3.6 Lessons from the pilot study

This section explores what can be extrapolated from this case study of an evaluation within a programme in three ways: the wider implications for evaluation as an aspect of the English language curriculum; the methodological implications for the second case study in this research; and the implications for the conceptual focussing to refine the research questions, and guide data collection and analysis in the later chapters of this thesis.

a) The wider implications

This holistic account of a programme illustrates the extent to which evaluation is interwoven into the fabric of the programme, linking the institutional evaluation policy; formal aims and objectives; teaching and learning styles and strategies; classroom interactions; and learning materials and other resources. Some tentative conclusions and issues for consideration in the wider ELT field are set out in Kiely (1998). For reasons of space these are not repeated here where the emphasis is on the implications of this study for the development of the thesis as a whole.

b) The study of evaluation - methodological implications

The nature of the observations and emergent hypotheses discussed in 3.4 above suggest that the evaluation policy and implementation are problematic in terms of both curriculum development and quality enhancement on the one hand, and the roles of

teachers and students on the other. These issues relate to classroom cultures and institutional practices, and are evident in the dense fabric of pedagogic and interactional routines in the classroom. Thus, the research strategy used here - a qualitative, ethnographic approach - seems appropriate, provided that the data issues outlined below are developed.

Three methodological points relating to the data demands of hypothesis exploration implicit in the research strategy emerge from this study. First, there is a need for more complete process data on the programme, of the different programme constructs - the institution's specification; the teachers' plan; and what happens in the classroom. As discussed in 2.5 above, we can make assumptions about what an ELT programme involves. It is clear from this pilot study, that we cannot understand the role and contribution of evaluation of the programme without a parallel, detailed study of what the programme involves.

Second, key interactions in the classroom need to be more fully documented than is possible with a reliance on observation notes. Understanding the student contribution to the evaluation in the small group discussion stage requires a comprehensive account of the interaction, which a video or audio transcript could provide. Thus, the hypotheses emerging in the study regarding the management of the interaction in the group discussion - the consensus seeking role played by Carrie, the dominance sought by Abe - could be developed and confirmed.

Third, more attitudinal data is required to guide and triangulate the interpretations of observation data. The notion of *agenda*, as used above, suggests that the participants have ideas about themselves and their role which underpin actions and utterances.

Millie's comments on my observation notes illustrate the need to capture these:

Millie: To some extent they were not a bad picture of students' reactions to things, and the interaction that was going on. But there is still - it misses out the students' interaction with the texts, and the tasks they're actually doing, and what effect that has on them. That's very difficult

to measure, but that is a big part of the lesson. They are doing an activity - you can comment on how interested they seem to be in it, how they go about it, but you do not necessarily look at what they were producing, or what the materials themselves were trying to achieve.

Interview with Millie CS1:12

Thus, teachers have rationales for plans and decisions they make, and students equally have a schematic frame which informs their interpretations, judgements and behaviours. The main way in which this body of data can be gathered is through interviews with teachers and students across the programme, which are largely unstructured, but cued to focus on aspects of and issues within the programme.

The issues, together with a discussion of ontological and epistemological traditions in social science, education and ELT research, inform the elaboration of the research strategy and design for the main case study in Chapter 4 below.

c) The study of evaluation - conceptual foci

The study of evaluation above follows the chronological structure of the process - an approach used by Weir and Roberts (1994 - see Figure 2.3 above), and the studies of evaluation in Alderson and Beretta (1992), Rea-Dickins and Lwaitama (1993), and Rea-Dickins and Germaine (1998). Within the narrative, however, there are issues which inform on the nature of this form of evaluation. These can be related to the research questions outlined for this study as follows:

RQ1 How do teachers manage the different agendas - their own personal and professional interests, students' concerns, institutional requirements - involved in carrying out evaluations?

One answer to this question is that the teachers focus on the interaction in the classroom - they keep this on track through a clear personal and professional sense of direction for the programme, and minimal attention to the students' concerns and institutional requirements (see Millie's comments page 69 above). This research

Chapter 3

question thus represents the set of conceptual foci to be explored in relation to evaluation in the second case study. Chapter 5 explores four of these:

Research questions	
1.1 The nature of the programme evaluated	To what extent are there different programme constructs: the formal specifications of the institution; the teachers' plan (what the 'persuasion' focusses on); and the series of classroom activities? How does the evaluation relate to these?
1.2 The nature of the evaluation events	What are the conceptual features of the evaluation procedures, with special reference to NGT? To what extent do these reflect the programme (validity) and to what extent are issues not considered (threats to validity)?
1.3 Characterisation of learning within the programme	To what extent is there a single learning agenda? To what extent are students aware of this? To what extent is it developed through the evaluation?
1.4 The consequences of the evaluation	What happens with the programme after the NGT procedure, that can be evidentially linked to the evaluation discussion? To what extent does this represent an improvement to the programme? How does it relate to the construct (or different constructs) of the programme? How does it relate to future programmes, and to teacher development?

Chapter 6 examines the role of the institution in the evaluation, and more widely, in the programme:

1.5 The role of the institution	In what ways can the evaluation be said to enhance and assure quality? How does this function affect the teacher's role? To what extent does the evaluation raise questions of ownership (of the programme and the evaluation)?
------------------------------------	---

Chapter 7 explores how the programme translates into a learning experience for one particular student, Laure, and how her participation in the evaluation relates to this.

1.6 The experience of the student	What is the student experience of the programme (and how does this experience relate to different programme constructs), and the evaluation? How do the evaluation functions - needs analysis, negotiation of programme emphases, etc. benefit the student?
--------------------------------------	---

Thus, RQ1 is seen as an umbrella question generating the six specific foci (1.1 - 1.6) to be explored in the context of the main study in this thesis.

RQ2 To what extent is involving students in evaluation a separate process from teaching them?

Two functions of the evaluation studied here are essentially pedagogic: the awareness-raising in relation to learning needs, and the negotiation of emphases for the rest of the programme. These evaluation issues are explored further as aspects of RQs 1.1, 1.2 and 1.5 outlined above.

RQ3 How do teachers use the time available to carry out the evaluation and deal with other pedagogic demands?

This issue merges in many ways with the question of teacher role discussed above. However, time in class is constructed as limited, and therefore valuable, a view shared by teachers and students. Time in this sense is explored in relation to RQs 1.2 and 1.5 above.

RQ4 To what extent are the evaluation findings valid and reliable?

The discussion of the literature in Chapter 2 above focussed on evaluation as an inquiry process, informing various stakeholders about various aspects of the programme. As a process of inquiry, the key value must be validity⁸. It is in this sense that the concept of validity is important for the analysis chapters of this study. RQs 1.1-1.6 above illustrate the different ways validity needs to be explored when looking at evaluation as a dimension of the curriculum. Chapter 5 explores how four aspects of the evaluation process within the programme (RQs 1.1-1.4 above) cohere with the purposes and principles of evaluation. Chapters 6 and 7 (RQs 1.5-1.6) examine the

same theme in macro and micro contexts - the institution and to the individual student.

RQ5 To what extent do evaluations focus on problems and explain them such that improvements are brought about?

Problem-solving and improvement have become fundamental components of evaluation theory and practice (see Chapter 2 above). These aspects of evaluation are, like validity, are strands in the complex weave that programme evaluation by teachers and are explored as part of the refined RQs 1.1-1.6.

RQ6 To what extent is the experience of others - relevant literature - used in investigating problems and devising solutions?

The use of the literature in the published evaluation studies surveyed in Chapter 2 corresponds to conventions in research and other intellectual fields. In the case study discussed in this chapter, there is little obvious reference to the evaluations or findings of other. Processes instead, are guided by institutional and departmental policy, and the pedagogic demands of the classroom - issues examined in RQs 1.1-1.6. The more general issue - the technical and bureaucratic, rather than intellectual, nature of this form of evaluation - is revisited in Chapter 8 below.

RQ7 To what extent do evaluations influence the pedagogy in terms of classroom activities, materials, assessment procedures, etc.

This question, in many ways a dimension of the 'improvement' issue discussed above, is taken further in the following chapters in two ways. First, the analysis of RQ 1.4, in relation to vocabulary (see 3.5b and 3.5c above) traces how a pedagogic strategy is explored and changed as part of the evaluation process within the programme. Second, RQ1.6 explores how the teaching strategies and evaluation processes shape the programme experience of one particular student.

3.7 Summary

This chapter describes the institutional context of an EAP programme, the evaluation policy which has evaluation by teachers using the NGT procedure at its centre, and the implementation of this policy in one programme. The case study illustrates the relevance of research into evaluation processes for a better understanding of ELT programmes, especially in contexts characterised by a negotiated, learner-centred curriculum and a concern for quality management. This pilot investigation relates in three ways to the whole study. First, the data and discussion here foreshadow the problems of programme ownership, student role and contribution, and the links between programme content and evaluation findings, which are framed in refined research questions for further exploration in Chapter 5, 6 and 7. Second, this case study establishes some methodological requirements for data collection and analysis in the main study. These are discussed in Chapter 4. Third, this case study also serves as a body of data which can be integrated into the main analysis chapters, and informs the discussion in Chapter 8 on how this form of evaluation might be further researched and developed.

Chapter 4

Research Design and Methodology

4.1 Introduction

This chapter examines the methodological issues of this research study. First, it looks at research and evaluation as traditions and contexts of investigation. Second, the philosophical basis of the research approach is discussed. The task here is to set out the case for a qualitative, naturalistic and ethnographic research strategy, and illustrate how this approach coheres with the issues within the areas of EAP programme evaluation. Third, case study as a research tool is examined. Fourth, the research strategy, informed by ethnography, is described in detail in two sections: data gathering and data analysis. Finally, there is a discussion of the issues of validity and researcher role. This relates both the epistemological basis of my research study, and to the validity issues of evaluation processes in the classroom. A framework from Messick (1989;1996) is examined as a platform for exploration of the validity of the EAP programme evaluation in chapters 5, 6, and 7. The chapter concludes with a summary of the research design, linking the research questions elaborated at the end of Chapter 3 with the main analysis chapters.

4.2 Evaluation and Research

There are two principal ways of relating evaluation and research: first, they are seen as essentially different, and second, they are considered as practices sharing a common philosophical and methodological basis, but different, in terms of purpose and resourcing. A detailed account of the 'different' school is set out by Glass and Worthen (1971). This reflects a positivistic view of inquiry in the social sciences, where 'research' seeks to establish generic laws which predict social behaviours and processes, and evaluation is context-specific problem solving. Norris considers this account of the research and evaluation relationship as "asocial and ahistorical" (1990:100). He sets out a theory of evaluation of educational programmes, which has i) moved away from a

reliance on constructs in experimental psychology, ii) established a range of praxeological areas of educational and schooling processes as sub-disciplines, iii) assumed a responsibility to explain as well as describe, and iv) developed an epistemology which incorporates principles of ethics and autonomy (see 2.3 above). This view is very much in the Stenhouseian tradition of practice informing practice (Stenhouse 1975), and shares an epistemological base with the realistic evaluation set out by Pawson and Tilley (1997) in relation to programmes across the social sciences.

Mitchell (1989) sees research as the task of representing an educational process, and evaluation as one possible use for the account which emerges. Thus, evaluations she has carried out have research aims, for example:

- i) To document current instructional practice, and to explore the extent to which this was providing experiences likely to lead to L2 acquisition;
- ii) To document the attempts of “committed” teachers to implement a “communicative” approach;
- iii) To develop an operationalised model of communicative competence.

Mitchell 1989:196

The discussion in Chapter 2 above shows that such a shared theory-building and problem-solving task for evaluation has been a goal in the ELT field. However, it has remained a somewhat unfulfilled goal, partly because of research methodologies which prioritise theory over data: the initial assumptions of what the programmes involve obscure what the participants see as the programme, and ultimately the problem-solving is not successful. The Lawrence evaluation in Zambia (see 2.6c above), for example, contrasts with the more grounded approach to evaluation elaborated by Crossley and Bennett (1995).

The pilot study in Chapter 3 - a study of the nature and role of evaluation within a programme - illustrates on the one hand, the need to base evaluation on valid constructs underpinning the programme, and on the other, the complexity of these constructs. It shows the embeddedness of the evaluation, deriving from unspoken, ambivalent attitudes to evaluation on the part of the teachers, and unpredicted expectations of students. The methodology set out in this chapter for researching evaluation seeks to describe evaluation in its programme context, such that interacting phenomena of the programme and the evaluation are explicated and understood. The processes of programme evaluation - what the teachers do in their classrooms - and the research - my documenting of these processes - are held as separate in this study. However, ways in which the research methodology developed for this study might inform evaluation methodology for contexts such as this are explored in Chapter 8 below.

4.3 Research approach - issues of ontology and epistemology

Lincoln and Guba propose an alternative paradigm - a naturalistic rather than a rationalistic method of inquiry - in which the investigator avoids manipulating research outcomes.

(Lincoln and Guba 1985: Introduction)

This section looks at the paradigm debate in the social sciences generally, and in education and ELT in particular: a debate introduced in the context of ELT evaluation in 2.6 above. Although the paradigm debate in research in these fields is a process of finding the most appropriate epistemological base and empirical strategy for inquiry into a particular phenomenon, it has become in recent decades a political and ideological discourse, as evidenced by the statement above from the introduction to Lincoln and Guba (1985). Two points here merit mention, since the generally credible and coherent text is widely used in this study as a guide to the development of an appropriate research strategy. First, in identifying theirs as an alternative, they are ceding to the other the mantle of *mainstream* or *normal*. While this is less the case with British educational evaluation and research, where various forms of naturalistic inquiry have occupied the central stage since the work of Stenhouse, it is still an issue

in ELT and applied linguistics, where “Paradigm wars linger on” (Edge and Richards 1998:335). Second, when they state the other school ‘manipulates research outcomes’, they assume for themselves the moral high ground, with a monopoly on trust and credibility. A debate cannot be furthered when one side refuses to accept the legitimacy of the other school: a resolution can only come through a recognition of the core values, and demonstration of how these can be met. The task here is to examine the ontological and epistemological bases for dissent, and while accepting the relevance to research of each side, seek the most appropriate generator of a research strategy for this particular task.

Cohen and Manion (1995) set out four assumptions underlying inquiry into human activity where the subjectivist and objectivist approaches differ:

Subjectivist	Assumptions regarding	Objectivist
<i>No</i> Nominalism	<i>Is social reality external to the individual?</i> ← Ontology →	<i>Yes</i> Realism
<i>No</i> Anti-positivism	<i>Is knowledge hard, tangible facts?</i> ← Epistemology →	<i>Yes</i> Positivism
<i>No</i> Voluntarism	<i>Do environmental factors determine human actions?</i> ← Human nature →	<i>Yes</i> Determinism
<i>No</i> Ideographic	<i>Is the purpose of research to discover general laws?</i> ← Methodology →	<i>Yes</i> Nomothetic

Figure 4.1 Assumptions underlying inquiry in the social sciences (after Cohen and Manion 1994:9)

This discussion in Cohen and Manion is based on Burrell and Morgan (1979), and the approach is essentially setting out the objectivist, scientific approach, and an alternative - where the epistemology is simply labelled anti-positivism.

Guba and Lincoln (1985) explore this anti-positivist alternative more fully, and relate it, not only to traditional positivism as in the natural sciences, but also to post-positivism, where the elements of interpretation, and non-empirical theory development are acknowledged.

Lincoln and Guba state that the first step in a process of inquiry is making a decision about the axiomatic system to proceed from, and that this decision is “best made by testing the fit between each system and the case”¹. Thus, the researcher uses existing understandings about an area of activity to select the research approach. The starting point for both schemes outlined above is the ontology: is the reality such that a conceptual framework can be elaborated which links theoretical definitions with demonstrable and quantifiable measures in the real world, and which can produce a cause and effect explanation for the phenomenon? Or, is the reality a complex of different interpretations in a given conceptual area, different constructions based on individual experiences, which the research seeks to understand? In the case of the Cohen and Manion framework (Figure 4.1), it is clear how the research process unfolds from a “Yes” answer to the first question: If human activity can be broken down into knowable component parts, then a rationalist, objective approach can provide explanations: accounting for what is done in all cases and contexts. Thus, the research reveals laws which determine behaviour. The method for generating this epistemology is nomothetic, paralleling the process of research in the natural sciences.

An early critique of this approach to understanding social processes is Lewin's (1946) proposal for action research (see 2.8a, page 41 above), which prioritises context over grand theory, the local and idiographic over the general and nomothetic. This seminal critique has proved enduring. It engendered a post-modernist, post-positivist body of evaluation and research, especially in education, which explores the local as something interesting in itself, and as the most relevant means of understanding the problems which prevail, and taking action to resolve them. The relationship between the physical and social sciences posited by Lewin is especially instructive: it

illustrates the importance of ‘human nature’ assumptions outlined by Cohen and Manion (Figure 4.1 above), and the epistemological axioms of Lincoln and Guba (Figure 4.2). It obliges us to attend to the voluntarism of human actions, and the multi-faceted, multi-functional nature of roles, agendas, and utterances.

Naturalistic paradigm	<i>Axioms about</i>	Positivistic paradigm
Realities are multiple, constructed and holistic	<i>The nature of reality (Ontology)</i>	Reality is single, tangible, and fragmentable
Knower and known are interactive, inseparable	<i>The relationship between knower and unknown (Epistemology)</i>	Knower and known are independent, a dualism
Only time and context bound working hypotheses (ideographic statements) are possible	<i>The possibility of generalization (Epistemology)</i>	Time- and context-free generalizations (nomothetic statements) are possible
All entities are in a state of mutual simultaneous shaping, so that it is impossible to distinguish causes from effects	<i>The possibility of causal linkages (Epistemology)</i>	There are real causes, temporally precedent to or simultaneous to their effects
Inquiry is value-bound	<i>The role of values</i>	Inquiry is value-free

Figure 4.2 Naturalism and positivism (after Lincoln and Guba 1985:37)

The naturalistic paradigm celebrates the contribution of values: they provide an ethical base for both the conduct of research, and the validity of interpretations and findings. Lincoln and Guba set out five corollaries that arise from a values-bound approach to research. Figure 4.3 shows how these values relate to the task of this thesis: research into programme evaluation by teachers.

<i>Corollaries of a values-bound approach to inquiry (Lincoln and Guba 1985:38)</i>	<i>Implications for current study - this corollary means:</i>
1. Inquiries are influenced by <i>inquirer</i> values as expressed in the choice of problem, evaluand, or policy option, and in the framing, bounding, and focussing of the problem, evaluand, or policy option.	<ul style="list-style-type: none"> that the issue is selected by me, the researcher (see Chapter 1), and that there are reasons for this which need to be taken as part of the research as a whole; in particular, the episodes analysed in Chapters 5, 6 and 7 derive from the data, and also from my initial interpretations of what these data might mean.
2. Inquiries are influenced by the choice of the <i>paradigm</i> that guides the investigation into the problem.	<ul style="list-style-type: none"> that the naturalistic approach has been opted for by me, and that this is indicative of how the problem is perceived prior to data: i.e. while the programme and evaluation derive from an outline plan, what is involved in the process of each cannot be predicted, but can be understood in relation to co-occurring phenomena.
3. Inquiry is influenced by the choice of the <i>substantive theory</i> used to guide the collection and analysis of data and in the interpretation of findings	<ul style="list-style-type: none"> that the approach to understanding the problem - constructivist and ethnographic - has been opted for, in order to facilitate the elaboration of differing agendas and interpretation of phenomena, which can be related to often conflicting features of a single context and community of practice (see 7.5 below).
4. Inquiry is influenced by the values that inhere in the <i>context</i> .	<ul style="list-style-type: none"> that research design decisions are taken to conform to the ethos of the institution and practices being investigated, and to ethical principles in representing teachers, students and the institution where the research is conducted. These principles involve doing no harm; not being unduly intrusive, particularly into demands on time of informants; adding value where possible, in terms of providing requested feedback by teachers and students (see 4.8 below).
5. With respect to corollaries 1-4 above, inquiry is either <i>value-resonant</i> (reinforcing or congruent) or <i>value-dissonant</i> (conflicting).	<ul style="list-style-type: none"> that the research sees all these positions as congruent. As foreshadowed in Chapter 3, however, there is an element of value-dissonance in evaluation policy and practice (see 6.4 below).

Figure 4.3: Summary of values underpinning research design

This approach to the research led to my selecting a case study approach as a way of defining the context and research subject, and a broadly ethnographic method of data collection and analysis (see 3.4 above). The next section examines the literature on these research tools, and relates them to the task in this thesis.

4.4 Research approach 2 - Case study

Stake (1995) identifies three kinds of case study: intrinsic case study, instrumental case study, and collective case study. An intrinsic case is studied because it is in itself interesting; an instrumental case is studied in order to inform on a class of phenomena of which the case is one; a collective case study is one where a set of phenomena are studied in order to inform on how they interact and relate to each other. An example of the first type is a biography: it is studied because the subject is interesting, and there is a potential for wider learning from understanding the nature and weave of the strands that constitute a life. Holliday (1996) characterises the notion of *an ethnography* in this way. There is no real prediction potential here: one cannot, and does not expect to learn how another life might develop from the study of one. Nor is there explanation in terms of precise cause-effect conclusions in relation to the events recounted in the biography. For these reasons, intrinsic case study has had little impact on the organisation of research within social science. The more orthodox view (Yin 1984; Nunan 1990; Cohen and Manion 1995) of case study is what Stake labels instrumental or collective: the task here is to study a case in order to generalise to a wider set of cases; to inform in depth on one of a range of cases being studied; or to identify through close scrutiny of one case, a framework of factors or issues for examining a class of cases. Macdonald and Walker (1975) provide a characterisation of intrinsic case study particularly resonant for this study: case study has the potential "to reveal properties of the class to which the instant belongs" (1975:4) where that instant is embedded in a "complex set of politically sensitive relationships" (1975:7), and where "at all levels of the system what people think they are doing, what they say they are doing, what they appear to others to be doing, and what they are in fact doing, may be sources of considerable discrepancy" (1975:8).

It is significant that key proponents of case study in educational research are also evaluation theorists (Macdonald 1976; Simons 1981; Stake 1995; 1996). Stake (1995) considers all evaluation as case study, (although not all case studies as evaluations). They are holistic ways of understanding what the various planned elements of a programme mean for the operation of the system and the lives of those participating.

A case is a “bounded system” (Smith 1978). For both Smith and Stake, boundedness has, in ethnographic terms, an *emic* quality; it is grounded in the experience and perceptions of those working within the system. This view has implicit tensions when applied to the focus of this study: programme evaluation by teachers. The evaluation procedure may be bounded as a policy of the institution, or as an episode in the classroom. However, for teachers and students the boundaries between evaluation and teaching, individual learning experiences, materials, learning needs, etc. may be unclear or even unreal. For example, the awareness-raising efforts of Millie and Anna in CS1 (see 3.5a above) with regard to independent learning might be part of a teaching strategy rather than an evaluation process. For teachers and students, the programme, or module, may be a case: an integral conceptual unit, with a clearly bounded set of features. However, for Millie (see 3.5c above) the key conceptual unit is her developing pedagogic approach, with individual programmes as a minor detail in this bigger picture. The problem is one of perception of boundedness: whether it is an *etic* concept, constructed by the researcher prior to fieldwork, or an *emic* one, conceptual strands teased from the complex weave of programme phenomena, deriving their unity from coherence and the potential for explanation. This latter form of boundedness, akin metaphorically to the gravitational pull of the core in a planetary system, rather than high fences on the perimeter of a domain, is exemplified, for example, in the case of the off-campus examination (6.2 below) and the study of Laure in Chapter 7.

Where a case is a period of time - a function of the calendar rather than a conceptual system, it is considered an 'episode', a borrowing from the field of ethnomethodology. Cohen and Manion, paraphrasing Harré, a realist philosopher, describe an episode as

any coherent fragment of social life. Being a natural division of life, an episode will have a recognisable beginning and end, and the sequence of actions that constitute it will have some meaning for the participants. Episodes may thus vary in duration, and reflect innumerable aspects of life.

Cohen and Manion 1995:205

Thus, episodes are the observable constituent parts of cases. The parameters of bounded systems are evident in behaviours and interactions. The analysis of a case is thus, a conceptual trailing of what may at first appear unrelated phenomena. For example, the issue of the off-campus examination in 6.2 below, profiles Anna's attitude towards non-teaching units of the institution, and informs on the evaluation of CS2.

The values-oriented approach to research outlined in Figure 4.3 above relates to case study in a number of specific ways. Figure 4.4 below sets these out using the six features generated by Adelman *et al* (1980) from a survey of the uses of case study in educational research.

Features of case study (Adelman <i>et al</i> 1980)	Interpretation for current study
a) Case study data is strong in reality but difficult to organise	'The case is a specific, a complex, functioning thing.' (Stake 1995:2). Representing these aspects of reality is facilitated by the grounded, ethnographic approach to data. The organisational problems are in the need to reduce the number of themes and issues presented and discussed, without being reductionist.
b) Case studies reflect the subtlety and complexity of the case in its own right, and allow generalisations	i) 'Case studies will often be the preferred method of research because they may be epistemologically in harmony with the reader's experience, and thus to that person a natural basis for generalisation' (Stake 1995:5). Thus, the onus is on me as researcher to achieve a)

to be made either about the instance, or from the instance to the class	above in relation to the case studied, and allow readers to perceive resonances and relevance for contexts with which they are familiar. ii) Case studies allow analysts to show how general regularities exist precisely when specific contextual circumstances are taken account of. When it is difficult to do this, then the theoretical formulation of the regularities needs some revision. In relation to evaluation theory, the study of practice in one particular context as in this study, may suggest that the theory needs to be adapted.
c) Case studies recognise the complexity and embeddedness of social truths, represent discrepancies or conflicts between participants, and also offer an alternative interpretation	Each case study is a description of a specific configuration of events in which some distinctive set of actors may have been involved in some defined situation at some particular point in time. The issues of the case - evaluation; the institutional policy; and implementation by teachers - are represented through the behaviours and voices of a distinctive set of actors, identified by names, with backgrounds, personalities and interests rather than uni-dimensional roles. Thus, both researcher and reader are appropriately equipped to identify patterns of reasoning and volition (Pawson and Tilley 1997) in a holistic context.
d) Case studies as products may form an archive of descriptive material sufficiently rich to allow subsequent reinterpretation	The concern throughout this study to develop a better description and theory of ELT programmes - to get beyond the War Stories and Romances (Swales 1989)(see 2.7 above). Each case constitutes a data set which can guide and be related to future cases, and which can be analysed to inform on, for example, issues of teaching methodology or materials. The construction here of a data set or archive is equally pertinent to the management orientation of this study: the programme cases are relevant to case study teaching methods such as those which characterise MBA courses.
e) Case studies are a step to action - they begin in a world of action and can contribute to it	While this feature is less relevant than the others to this study - the context of the research being a PhD thesis - the underlying perception of a professional and management problem in relation to this form of evaluation practice supports its potential for ameliorative action. This potential is briefly explored in Chapter 8. It is, however, more immediately translated into action, in my (the researcher's) professional action. (See discussion of reflexivity in 4.5 and 7.5 below)
f) Case study reports are accessible to a range of audiences.	Stake sees case study 'as telling the story' (1995). The episode approach outlined above illustrates the element of narrative in representing phenomena within the programme and evaluation. This derives naturally from the representation of abstract concepts and phenomena through the voices of the participants in the programme and evaluation.

Figure 4.4 Features of case study approach related to this study

Case study is not a method, but rather a framework for a study, a means of understanding the parameters of what is to be studied (Stake 1995). This concept of case study specifies what the data are to be about, and how they relate to the research questions (see 4.10 below), but not how they are to be gathered and analysed. The following section discusses the role of ethnography in the collection, construction and analysis of data for this study.

4.5 Research approach 3 - Ethnography

The method of data gathering and analysis used to operationalise the naturalistic inquiry approach of Lincoln and Guba (1985) is ethnography. Bloom and Green (quoted in Roberts 1998) distinguish between three characterisations of ethnography in educational inquiry. *Doing ethnography* involves “a broad, in-depth, and long-term study of a social or cultural group, meeting the criteria for doing ethnography as framed within a discipline or field” (1998:56) This characterisation shares much with Stake’s (1995) intrinsic case study, and Holliday’s (1996) notion of *an ethnography*. *An ethnographic perspective* uses the theory and inquiry practices from anthropology and sociology to do a more focussed, less comprehensive ethnography. Using *ethnographic tools* involves fieldwork methods and techniques which may or may not be guided by cultural or social theories. This framework is useful for understanding the potential of ethnography to the field of language learning and teaching on the one hand, and the concerns of evaluation on the other. The first characterisation has limited application to the field, given the difficulty with defining a social or cultural group appropriate for a broad, in-depth and long-term study, and also the issue of whether it would be useful to study such a group for its own sake. An ethnographic perspective is represented in Van Lier (1988), Roberts *et al* (1992), and Holliday (1994), while using *ethnographic tools* is evident in a range of ELT and evaluation studies which seek to represent emergent issues from a practitioner perspective (e.g. Woods 1997; Kiely 1996; Dubin and Wong 1989). This study envisages the use of ethnographic tools to enhance the quality of observation and interview data. (However, see 7.5 below).

Roberts *et al* (1992) sets out eight reasons for the use of ethnography in the study of language education programmes. Figure 4.5 relates these to the key elements of the Lincoln and Guba framework for naturalistic inquiry.

Reasons for ethnography (Roberts <i>et al</i> 1992:180)	Interface with naturalistic inquiry (Lincoln and Guba 1985:188 - See also 3.4 above)
i) a respect for those studies based on trying to understand people's world as they see it	Human instrument Negotiated outcomes
ii) a focus on the common-sense, the ordinary and the everyday	Natural setting Human instrument Grounded theory
iii) the use of common-sense methods of investigation to capture how people make sense of the world in a routine way	Human instrument Qualitative methods Tacit knowledge Inductive data analysis
iv) a belief that people's behaviour can be studied only in context	Natural setting
v) flexibility: a resistance to pre-conceived ideas about what will be discovered and a methodology which starts with observation in the field rather than designing a range of fixed methods before the fieldwork begins;	Human instrument Emergent design Inductive analysis Grounded theory
vi) a focus on the detail of people's everyday lives and an acknowledgement of the complexity of social interaction	Human instrument Qualitative methods Purposive sampling
vii) an attempt to see the strange in the ordinary so that no aspect of the routine and detail of everyday lives is overlooked	Human instrument Ideographic interpretation

viii)	the use of a range of elicitation techniques and analysis: <ul style="list-style-type: none"> - observing/listening - audio- and video-recording - interviewing - documenting the physical environment - documenting patterns of movement and interaction 	Qualitative methods Purposive sampling
-------	--	---

Figure 4.5 Reasons for using ethnography as form of naturalistic inquiry

These reasons reflect the values orientation of the study (see Figure 4.3 above), and provide a methodology to develop data sets consistent with the case study approach set out in Figure 4.4. The data gathering and analysis procedures generated by this broad research strategy are set out in the following sections.

4.6 Research strategy 1 - procedures and instruments

The data sets for the whole study - Case Study 1 (CS1), the pilot study discussed in Chapter 3, and Case Study 2 (CS2) - are summarised in Figure 4.6. The remainder of this section sets out briefly the rationale for each data-gathering technique, and some issues which arose in the data gathering process. In keeping with the reflexivity implicit in naturalistic and ethnographic methods, issues to do with the gathering of data are considered an integral part of their meaning (Lincoln and Guba 1985:327). Thus, the analysis chapters of this include discussion of data issues and problems - see for example 7.5 below).

Data source (See Appendix 1 for volume of these data sets)	CS1	CS2
Classroom observation	21 hours - all of Millie's classes (18 hours), and approximately 3 hours taught by Anna or team taught	36 hours - all 12 sessions of the programme, taught by Anna
Interviews with teacher, based on interview schedule (Appendix 3)	3 interviews with Millie, 1 based in initial questionnaire, and 2 on evaluation session in CS1	6 interviews with Anna (see Appendix 4)

Interviews with students based on interview schedule (Appendix 3)	-	23 Interviews with 14 students, including 6 pair interviews, and 5 interviews with Laure (see Appendix 4)
Interview with coordinator based on interview schedule (Appendix 3)	-	1 interview with coordinator for EAP programmes, based on interview schedule (see Appendix 4)
Audio recording of evaluation group discussion	-	Recording and transcript of one group of six students evaluating the module as part of the NGT technique
Questionnaire study of EAP teachers	Questionnaire on professional history and use of evaluation based on Sanders (1992), returned by 4 teachers, including Millie and Anna	
Programme management documents	Flyer promoting EAP modules Module guide Resources guide End of module marksheet	Flyer promoting EAP modules Module guide Resources guide Assessment criteria End of module marksheet
Classroom materials and learning tasks	All classroom materials used in sessions observed	All classroom materials (See analysis of materials Appendix 5)
Written account by students on their background and learning needs	-	Essay titled 'Me' set by Anna. Copies of 16 students' work (See Appendix 7 for two examples)
Written evaluation by students	-	Written task set by me after mid-module evaluation Copies of 10 students' work (See Appendix 7 for two examples)

Figure 4.6 Summary of data collected on two programmes (CS1 & CS2)

a) Classroom observation

A key influence for the use of an ethnographic observation approach is Mitchell's account of the limitations on systematic approaches, summarised in 2.6 c) above. Since evaluation (especially for management purposes) and teaching tend to be seen as representing a conflict of cultures (see Chapter 1 above and Chapter 6 below), ethnographic observation is an appropriate tool for documenting both cultures and conflict. The technique adopted for CS1 (see 3.4 above) was therefore, used again for CS2 to document:

- i) the physical space of the classroom - number of students; seating arrangements; movement within the classroom; phenomena in the vicinity perceivable in the classroom;
- ii) the temporal space of each session: what happened when and how long it endured;
- iii) acts of teaching and learning - what the teacher did (including notes on what was said), what the students² did, the interactions in the classroom, and brief descriptions of the materials.

A goal in the construction of these fieldnotes is documentation of all that I - the human instrument - perceived. Some perceptions are evidenced physically, for example:

9.10 Sao and Rata come in - Rata sits in the same seat as last week, Sao next to her. Anna hesitates while they do that then continues.

Classroom notes CS2:3

Others involve interpretation on my part, signalled by the use of *seem* or *like*, for example, when Anna challenges Eli for stating that the Turks are the cause of ethnic problems in Cyprus:

Anna: Perhaps you could start with that? Also you are giving an insider view, as if your position was the natural one, a Greek, one of us, as opposed to them. Eli takes this up, making the point that the Greeks do not have just one religion. I am not clear whether he has understood the point - he has understood one aspect of it - that he as a Greek is siding with them, in stating that they, the Greeks, are pluralistic and tolerant. Anna seems to be aware of how delicate all this could be and says:

Anna: There are complexities.

Classroom notes CS2:3

The notes document the difficulty in deciding whether this represents a developing interaction, or a form of communication breakdown. And, in addition, they seek to represent how Anna is interpreting the situation, thus suggesting an analysis which

incorporates her closing utterance. This illustrates how the process of data gathering is in part a process of analysis. As such it requires triangulation and confirmation. The technique for this was to leave a space in the typed up classroom notes and ask Anna to comment. Usually she responded verbally, either informal discussion (which I would then headnote and include at the end of the classroom notes) or in interview. An example of triangulation with interview, in relation to Eli, illustrates how the technique was successful both in documenting the classroom process, and in verifying the emerging account.

Anna: [...] I want to help them to understand the difference between an insider perspective, and let's say an academic perspective. An insider perspective is like the Greek guy was doing, he was saying: this is the situation in this ethnic conflict, they have done this, they have done that, and it is ridiculous, blah blah blah. Fine, as a kind of polemical, sort of, talk with your friends, kind of thing

Interview with Anna CS2:4

The classroom observation data thus provide a thick description of the actions and interactions in the classroom, rich enough to facilitate analysis of emergent themes, but containing also a certain element of analysis as the real world of the classroom is transformed into text. The notes serve to generate hypotheses which the other data (see Figure 4.6) inform on further, confirm or refine.

b) Interviews

The account of CS1 relied on classroom observation, the interviews with Millie providing information on Millie as a teacher, and her thinking at the end of CS1. In addition to the need for a data source to triangulate the classroom notes (See 4.6a above), I wanted to people the narrative, to develop an account of the programme and the evaluation in the voices of the participants. As Figure 4.6 and Appendix 4 illustrate, interviews with the teacher and students represent a significant diachronic perspective of thinking across the programme. The interview schedules designed to frame interviews are attached as Appendix 3.

The ethnographic orientation of the research strategy determined both the ethical basis for the interviews and the procedure developed as the interviews were carried out. Woods (1986) sets out three attributes of the social situation where the ethnographic observation and interviewing is taking place. First, a relationship based on *trust* between the interviewer and interviewee is required. Second, the researcher needs a level of *curiosity* which constitutes a drive to explore and understand the phenomena prevailing in the situation. Third, the interviews should have a *naturalness*, such that the data represents as far as possible the perspective of the interviewee, and is not determined by leading questions and the particular line of inquiry of the interviewer. Within ethnography, mandated data is a threat to validity, as it relates to an external *etic* construction of the phenomenon at the expense of an internal *emic* perspective, the goal of an ethnographic approach.

In my research context, the trust factor was not a problem: I had worked and was at that time working closely with Anna on other projects, and the students, some of whom were known to me from other curricular contexts seemed without exception eager to cooperate. The curiosity factor was there but not so easy to manage: following up every issue was not possible due to the limited time of others, especially Anna. The third factor, naturalness, also presents a tension. While interviews should resemble a conversation, there is a shared expectation in this situation that the interviewer will ask questions and set topics for the interviewee to elaborate on. In this study, the technique used was to set out the materials from the sessions on the table, and ask the interviewees to talk about them. Some students (for example, Laure - see Chapter 7 below) talked freely and expansively in these situations. Others, such as Jamie who was at the lower end of the proficiency in English scale, needed a more directed approach, which seeks to recount experience in a practical concrete way:

- I: Were you able to complete the tasks in the time you had?
J: Yes, I think so, yes, the thing is, almost always I need a dictionary, to help me.
I: Did you use a dictionary today?

- J: No, because I was with the teacher, but almost always I need a dictionary, for writing the essay.
- I: And for reading, reading something like this?
- J: Yes, yes, and because for me it is easier to speak and to understand, but when I read something it is different, sometimes I don't have the idea that how it is, the spelling of that word.
- I: I see, so if we can look just at this section here, can you tell me about your use of a dictionary, what words here in this paragraph, would you have liked to use a dictionary?
- J: I think firstly, *muffled*, *purr*, that's it.
- I: Nothing else? That's OK.³

Interview with Jamie CS2:11

In addition to the programme participants, I also interviewed the coordinator for all the English for Academic Purposes programmes. The purpose here was to gain purchase on the role of the department and wider institution: to what extent do influences other than the teacher determine what is done? And to what extent do these shape the evaluation of the programme? The questions / cues (see Appendix 3) explore what normally happens in the 'coordination' of these programmes, and the extent to which there is an institutional view of what students' needs are, and the kinds of materials and activities which best meet these. This interview also had an important validating function for my own position as 'human instrument' and 'insider': I was mindful of the extent to which I was a developer and user of the evaluation and quality assurance systems in place, and felt that there was a risk in relying on my assumptions in exploring the institutional contribution to a given programme.

c) *Recording of one group of students holding evaluation discussion*

In the pilot study reported in Chapter 3, a key piece of data was the discussion by a group of students as part of the mid-programme evaluation (see Figure 3.2 above). There, the data were constituted from notes. A concern in the main study was to improve the validity of the account of the classroom discussion which students had as a prelude to the plenary evaluation discussion in Week 5. I set up the recording with a minimum of intrusion: with prior agreement from Anna and the students concerned, I

recorded the evaluation discussion of one group of six students. Although the quality of the recording was variable - not all six students were sufficiently close to the microphone, and a number of times two or more conversations were going on simultaneously - I was able to transcribe it and get copies to four of the students involved to check and augment. Two students - Laure and Arnie - got together and added various notes to the transcribed. These did not correspond to inaudible sections of text in the recording, but tended to represent the substance of their comment. The recording provides a key piece of data for understanding the perspectives of individual students, the nature of the evaluative discussion, and the process whereby individual perspectives develop into the overall account in the plenary feedback session, or do not (see 5.3 below).

d) Documentation

The programme-related documents (see Figure 4.6 above) complete the data set for the study. The module management documents provide an account of plans and intentions, both on the part of the teacher and the institution. They thus facilitate an analysis of how the programme and evaluation in practice compare with the outline construct. The classroom materials and tasks represent both the classroom pedagogy, and a means for the participants to recall their experience in the classroom (see Appendix 5). They were collected and coded systematically over the twelve weeks of the programme. Each activity handout was coded: date/month/number in sequence for the day, for example, the handout for the third activity on 19 February is coded 19/2/3.

A somewhat different kind of documentation is the written work of students. I photocopied all samples of two tasks which students handed in to Anna. First, there is an essay entitled 'Me' which students wrote during the first session. Anna gave four reasons for this task: something to do while she was registering students, etc, to establish a focus on writing, to identify strengths and weaknesses, and to get to know the students as people. It was thus, a form of needs analysis similar to that described

in 3.5a above in CS1. The second piece of writing is a written evaluation of progress so far, designed by me - with Anna's agreement - to provide a channel whereby students could relate individual perspectives in addition to the general view of the group in the NGT procedure. Appendix 7 contains two examples of each: the 'Me' profile of Arnie, a student in the small group discussion (see 5.3 below) and Sue, a student who features in the vocabulary issue (see 5.5 below), and in the off-campus examination problem (see 6.2 below). The written evaluations are those of Sao, a student whose problems with the programme are analysed (see 5.3 below) and Laure, the student studied in Chapter 7.

4.7 Research Strategy - Data analysis and interpretation

The data in this approach are descriptive, determined by what is perceived rather than by any existing set of categories or codes. The main conceptual work in understanding a given process is at the data processing and analysis stage. Analysis is a particularly problematic area for research within a naturalistic framework. Miles (1979, quoted in Lincoln and Guba 1985) aptly characterises the analysis problem:

Qualitative data tend to overload the researcher badly at almost every point: the sheer range of phenomena to be observed, the recorded volume of notes, the time required for write-up, coding and analysis can be overwhelming. But the most serious and central difficulty in the use of qualitative data is that methods of analysis are not well formulated. For quantitative analysis there are a number of clear conventions the researcher can use. But the analyst faced with a bank of qualitative data has very few guidelines for protections against self-delusion, let alone the presentation of 'unreliable' or 'invalid' conclusions to scientific or policy-making audiences. How can we be sure that an 'earthy', 'undeniable' or 'serendipitous' finding is not, in fact, wrong?

Miles (1979) quoted in Guba and Lincoln 1985:354

Writers on ethnography (e.g. Hammersley 1992; Canagarajah 1996; Roberts 1998) also highlight the same need for attention to the methods of analysis in qualitative research and evaluation within an ethnographic framework. Within applied linguistics, the dominant method of analysis has come from techniques of discourse and conversational analysis (van Lier 1988; Roberts *et al* 1992; Hornberger 1995). Holliday (1994) and

Woods (1996) represent studies where the focus of analysis is pedagogy as a social process rather than as a linguistic construct. The present study seeks to develop an account using the same approach to the pedagogy - a programme made up of a series of episodes in the classroom (and beyond), not just a construct of the teacher, but a co-construction of the teacher, students, and institution (see 4.4 above). Thus, it seeks to construct a more social account, to build on the different perspectives of a range of the participants. The particular focus of this study is the evaluation of the EAP programme. The analysis therefore, has to find units of analysis which inform on the hypotheses relating to evaluation and related programme processes.

Lincoln and Guba present a useful discussion of data analysis approaches, based on a continuum elaborated by Goetz and LeCompte (1981) to represent the principal methods of processing and analysing ethnographic data. The approach here illustrates the degree of convergence between naturalistic inquiry and ethnography, evident especially in the domain of data analysis. The continuum set out in Figure 4.7 below involves four 'somewhat overlapping dimensions':

inductive / deductive
generative / verificatory
constructive / enumerative and
subjective / objective.

Arrayed along these continua are five analysis strategies:

- i) inductive analysis
- ii) constant comparison
- iii) typological analysis
- iv) enumerative systems, and
- v) standardised observational protocols.

The four dimensions here represent the central ontological and epistemological differences between the naturalist and positivist approaches to research discussed in 4.3 above. As should be evident from the discussion so far in this chapter, this study is positioned towards the inductive - generative - constructive - subjective pole. In other words, categories are not taken from existing theory, but developed from grounded analysis of classroom observation notes, triangulated with interview and other data. My research task is seen as the generation and linking of constructs which explain phenomena associated with classroom evaluation by teachers. This grounded approach is inherently subjective: it is developed from the perceptions and perspectives of the participants (including me, the researcher).

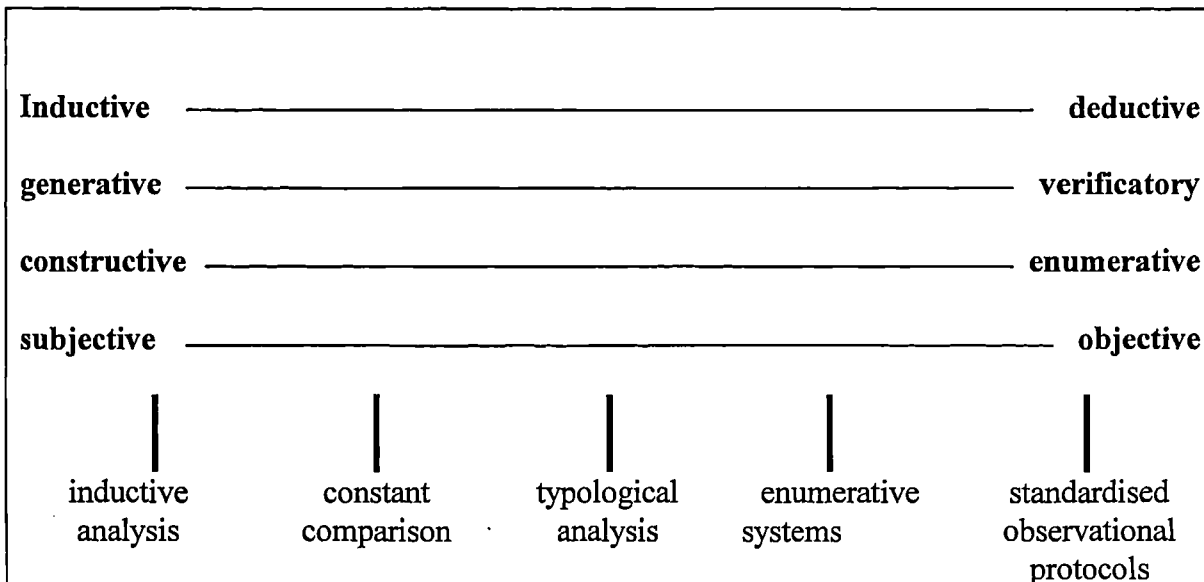


Figure 4.7 Strategies for analysis of ethnographic data (after Goetz and LeCompte 1981; Lincoln and Guba 1985)

Four of the Goetz and LeCompte techniques of analysis are used in this study (the exception being standardised observation protocols (see 2.6c and 4.6a above)). Inductive analysis is a technique where the data is scanned for evidence which questions or disconfirms a hypothesis which has emerged. The discussion in 6.2 below, tracing the development of the off-campus examination issue, and in 7.3 below, seeking to

understand the pattern in Laure's opinions about classroom materials, and her ambivalence regarding negotiating her essay title with Anna, illustrates this somewhat hermeneutic process of analysis.

Constant comparison is particularly relevant in this study where there was strong time and space structure to the observation: a class with the same teacher and student for three hours each week over twelve weeks. Comparing behaviours and interactions provides a means of identifying patterns, and especially, breaks in patterns. For example, an examination of vocabulary teaching activities analysed in 5.5 below illustrates how constant comparison provides both hypotheses and explanations for vocabulary teaching strategies, and informs on how these relate to the evaluation process and teacher development.

Typological analysis is a form of content analysis (Miles and Huberman 1984) which is naturalistic, where, as in this study, it is a form of organising, or constructing (Goetz and LeCompte 1981) the data, as a prelude to the other techniques of analysis. This is the approach used to in 5.2 to categorise the materials and teaching activities (see also Appendix 5). This analysis is then related to the context through a comparison with Anna's perspective on the pedagogy, and the institution's programme specification. The same construction of data provides a framework for understanding Laure's perspective on materials in 7.3.

Enumerative systems are those which prioritise quantified data, and as such play only a minor role in data analysis in this study. However, frequency counts and other quantifications assist in organising the data such that the times accorded to different activities in the classroom, information on student attendance, submission rates for written work, and assessment results can be set out in table form, and referred to easily when examining emergent hypotheses. Such data guide and augment other approaches to analysis; for example the use of the final assessment matrix affords an important

perspective on Laure as a student, and in analysing Anna's comments on strong and weak students.

The task in this section has been to forge a coherent research strategy for the study of evaluation within its curricular context, using the principles of naturalistic inquiry and the tools of ethnography. The rationale and procedural accounts represent a workplan, which, in keeping with the focus on process, on documenting the *emic* to develop grounded theory, itself evolves. Thus, issues of method are revisited in the main discussion chapters, especially in 7.5 where the issue of constructivism in ethnography is explored in relation to the analysis of a student case. This pervading attention to issues of method is a strategy for enhancing the quality of the empirical element of this study. The next section of this chapter relates the research strategy to the traditional dimensions of research quality - validity and reliability.

4.8 Validity as a research issue

Validity and reliability are especially complex issues in interpretive research. In the positivist research tradition, reliability is achieved through the size and random nature of the data set. Reliability thus amounts to replicability - the similarity of results arrived at by another researcher in a similar situation, and thus supports generalisation to wider populations. In interpretive research, where the unique, non-recurring nature of social interactions is essential to the validity of the data, reliability therefore cannot be seen as replicability, without compromising the validity - the documenting of the uniqueness of the interaction. Cohen and Manion (1995), discussing qualitative methods, conclude that there has to be a 'judicious compromise' in achieving both validity and reliability. Within the naturalistic inquiry of Lincoln and Guba (1985), validity and reliability are not considered as parallel pillars of research quality, but rather hierarchically-related aspects of a unified research quality: reliability is a requirement for, and therefore subservient to validity in all research.

Research instruments which are valid succeed in gathering data which accurately describe and represent the phenomena in question. Thus, the classroom observation strategy of my research has to document both the pedagogy and management of the programme in the classroom in order to be valid, and the interviews have to represent the perspectives on the programme of the various interviewees. Lincoln and Guba (1985) set out four criteria for validity, or trustworthiness, of naturalistic data. These are credibility, transferability, dependability, and confirmability and are related to the techniques used in this study in Figure 4.8.

Criterion Area	Techniques in naturalistic inquiry (Lincoln and Guba 1985:328)	Techniques applied in this study
Credibility	<ol style="list-style-type: none"> 1. Activities in the field that increase the probability of high credibility: prolonged engagement; persistent observation; and triangulation 2. Peer debriefing 3. Negative case analysis 4. Referential adequacy 5. Member checks (in process and terminal) 	<ul style="list-style-type: none"> • Observation over the duration and of all sessions of the EAP programme, with concurrent interviews with participants • Interviews and informal discussions with teachers and students • Identification of issues to raise in interviews • Interview strategy based on classroom materials; focus on <i>emic</i> • Passing classroom observation and interview data to teachers and students for comment within 48 hours of data collection.
Transferability	6. Thick description	<ul style="list-style-type: none"> • Ethnographic account of the classroom activities
Dependability	7. a) The dependability audit, including the audit trail	<ul style="list-style-type: none"> • Triangulation with other data sources such as interviews, and classroom materials
Confirmability	b) The confirmability trail, including the audit trail	
All of the above	8. The reflexive journal	<ul style="list-style-type: none"> • Headnotes, incorporated into classroom observation notes • Research journal, reflecting on researcher role

Figure 4.8 Techniques for establishing trustworthiness

These techniques, which centre on handling the data, combine with ethical and reflexive dimensions of my role as researcher to enhance overall research quality and thus, ensure the validity of the processes of data collection and analysis, and the research findings. Figure 4.9 summarises these techniques, informed by Lincoln and Guba's (1985) naturalistic inquiry.

Researcher role requirements	Instances in this study
Developing an ethical relationship 1) Being open about my interests and intentions 2) Not influencing unduly programme processes 3) Not inconveniencing the teachers or students in any way 4) Assisting them in any way I could	<ul style="list-style-type: none"> • introducing research project to teachers and in classrooms • unobtrusive observation in classroom • arranging interviews at time and place convenient to informants; not pressing informants for member-checked data • helping students with language-related and other queries; providing Anna with all interview data, at her request
Ongoing focus on research strategy 1) Documenting research impact 2) Ongoing development of research strategy	<ul style="list-style-type: none"> • Questions in interviews regarding research impact (See Appendix 3) • Reflection on what a data to theory approach means for data analysis (See 7.5 below)

Figure 4.9 Techniques for development of researcher role

The listing of guiding principles and techniques for implementation as in the tables above belies, however, the complexity of ensuring validity. First, there is always a tension between the time required to ensure accommodation of all perspectives, and the time available in the real world of a programme such as that which I studied.

Second, there may be a divide between the open, democratic research which I was committed to and the status-marked world of classrooms, where I was a teacher.

Third, validity has a strong process element: it is not so much a technologised procedure outlined and then followed carefully, as it is a negotiated, context-sensitive process, with the map becoming clear only as the journey is completed

4.9 Validity as an evaluation issue

Parallel to the issues shaping the validity of my research, is the question of validity of the evaluation process which is the focus of this research. Questions about whether programme evaluation by teachers is a good thing (see Chapters 1 and 2 above) raise issues of validity of policy and practice.

Four approaches to the concept in the evaluation literature illustrate how validity has largely been determined through reference to an epistemological tradition at a remove from the context and purposes of the evaluation.

- i) the internal and external validity of the experimental tradition (see 2.6a above);
- ii) the pragmatic perspective of the Joint Committee for Standards in Evaluation, which set out four criterial features evaluations should aspire to: utility, feasibility, propriety and accuracy;
- iii) the consequential validity of evaluators such as Stake (1995) and Patton (1995), who see the evaluation process as an integral part of the social programme in question, rather than a perspective on it from a distance which ensures objectivity and impartiality; and
- iv) epistemological validity, as set out by Pawson and Tilley (1997) for social programmes generally, and Mitchell (1989; 1990) in language teaching programmes, where the task of evaluation is to explain, to provide some account for how the strategies which make up the social programme are linked to the goals of the programme.

As a starting point for the analysis of the evaluation from a validity perspective, the framework outlined by Messick (1989) is used. Messick, working in the domain of assessment, takes an essentially stakeholder perspective: validity of the testing procedure is located in the experience of the tested, determined not only by the concept or design of the procedure but also its implementation and impact in a given educational context, and consequences beyond that context. He eschews the notion of distinct and independent separate types of validity: “it is not a type of validity, but the relation between the evidence and the inferences drawn that should determine the validation focus” (Messick 1989:16). This characterisation of validity is particularly useful for the form of programme evaluation described in this study where the process of data collection, analysis and interpretation all occur in real time. In CS1 (see 3.5c above) a limited amount of data was interpreted extensively to draw conclusions, a process which raises questions particularly about this relationship.

Messick identifies two facets of this unitary concept of validity: one is the ‘source of the justification’ whether based on evidence or consequence, and the other is ‘the function or the outcome’ based on the interpretation and the use of the data. Each of these facets has, thus, two separate bases - evidential and consequential. This analytical framework can be represented diagrammatically as follows:

<i>Facets of validity</i>	<i>Interpretation</i>	<i>Use</i>
Bases		
Evidential	Construct validity (i)	Construct validity + relevance / utility (ii)
Consequential	Value implications (iii)	Social consequences (iv)

Figure 4.10 Determining evaluation validity from Messick (1989:20)

We can relate this matrix to the evaluation process in the programme discussed here (mindful of Messick’s view that since validity must function as a unitary concept, the divisions in the matrix are for discussion purposes only and the reality is that there is overlap and interpenetration).

- i) Construct validity can be determined by looking at the degree of match between the programme and the evaluation procedure. This relates to a central question for the next chapter - To what extent is the evaluation an evaluation of the programme? In 5.2 below this question is explored through analysis of three programme constructs: as specified by the institution, as conceived and introduced by the teacher; and as the series of classroom activities.
- ii) Construct validity related to issues of relevance and use is examined in 5.3 below through the small group discussion data, an important phase of the NGT procedure. The focus is on data validity: the degree to which the students are knowingly participating in an evaluation procedure, and are basing their judgments on their actual experiences. The transcript and classroom observations data show how the NGT procedure is managed as an event, shapes the roles of teacher and students, and fulfils the pedagogic function of the evaluation. The discussion identifies threats to validity, and explores the reasons for these, and the extent to which they can be eliminated or reduced in this form of evaluation.
- iii) The value implications are for Messick an expression of how what might be considered a narrowly technical procedure within an educational discourse does in fact relate to the real world, both in terms of the aims of the educational programme and the teacher, and equally in terms of benefits for students. In 5.4 below, the values underpinning learning within the programme are explored through examination of one particular activity in the programme, which represents clearly the teacher's idea of what the learning within the programme should involve. In 7.2 below, the programme experience as informed by one student's values is explored. These analyses determine the degree to which values shape the programme, and inform the evaluation.
- iv) The social consequences as characterised by Messick, represent an area which has, until recently, received little attention in the testing literature. This is less the case

in evaluation where the question of impact has been explored from different perspectives. Stake (1995), for example, uses the term consequential validity to explore the impact an evaluation procedure has on a programme. In this case the focus is on the development function of evaluation, - the improvements to the programme and the teacher development - and the accountability function - quality assurance. In 5.5 below the data facilitate two different trails of one 'finding' or 'action point' of the evaluation: the classroom observation notes show how the pedagogic approach to dealing with vocabulary in texts changes. The interview data inform on the mechanisms of this change, and the impact it has on the teacher's beliefs and principles. In 6.3 below the contribution of the evaluation to the institutional quality management system is traced.

4.10 Summary of Research Design

This section summarises in tabular form the research design for this thesis which has emerged from the epistemological issues discussed in this chapter, the conceptual issues framed by the pilot study, and the implications of the Messick framework for valid evaluation practice. The three columns show the research questions (RQs 1.1-1.6 – see 3.6c above), the data sets used to explore them, and the context of analysis, both in terms of sections of the main analysis chapters, and the cases and episodes within the EAP programme.

Research Questions (RQ) (see 3.6c above)	Data sets used	Context of analysis (Numbers refer to section of main analysis chapters)
RQ1.1 1. The nature of the programme evaluated	<i>Classroom episodes</i> <ul style="list-style-type: none"> • Learning activities • Module introduction <i>Documentation</i> <ul style="list-style-type: none"> • Module specification • Learning materials 	5.2 Analysis of three different representations of the programme 5.3 The issues and findings emerging from the evaluation event 5.4 The teacher's perspectives on learning in the programme 6.2 The teacher – institution boundary

	<p><i>Interviews</i></p> <ul style="list-style-type: none"> • Teacher (Anna) • Students 	7.2 Analysis of the programme construct of one particular student
<p>RQ1.2</p> <p>The nature of the evaluation events</p>	<p><i>Classroom episodes</i></p> <ul style="list-style-type: none"> • Module introduction • Evaluation event: small group discussion plenary stage <p><i>Interviews</i></p> <ul style="list-style-type: none"> • Teacher (Anna) • Students <p><i>Documentation</i></p> <ul style="list-style-type: none"> • Evaluation policy • Module guide 	<p>5.3 Analysis of small-group discussion and plenary discussion stages of the NGT procedure</p> <p>6.3 The evaluation and quality assurance by the institution</p>
<p>RQ1.3</p> <p>Characterisation of learning within the programme</p>	<p><i>Classroom episodes</i></p> <ul style="list-style-type: none"> • Classroom activity <p><i>Interviews</i></p> <ul style="list-style-type: none"> • Teacher (Anna) • Students <p><i>Documentation</i></p> <ul style="list-style-type: none"> • Classroom materials 	<p>5.4 Analysis of one classroom activity which represents in concrete terms the teacher's construct of the module</p> <p>7.3 Analysis of one student's response to classroom activities</p>
<p>RQ1.4</p> <p>The consequences of the evaluation</p>	<p><i>Classroom episodes</i></p> <ul style="list-style-type: none"> • Evaluation event • Text based classroom activities Weeks 6-12 <p><i>Interviews</i></p> <ul style="list-style-type: none"> • Teacher (Anna) • Students 	<p>5.5 Analysis of one aspect of the pedagogy, raised in the evaluation and trailed through the data on the rest of the programme</p> <p>6.3 Analysis of how this trail relates to institutional programme development, and teacher led innovation within programmes.</p>
<p>RQ1.5</p> <p>The role of the institution</p>	<p><i>Classroom episodes</i></p> <ul style="list-style-type: none"> • Administrative issue - remote examination • Evaluation event 	<p>6.2 Analyses of the interface between the EAP programme and the institution</p> <p>6.3 Analyses of the purpose and</p>

	<i>Interviews</i> <ul style="list-style-type: none"> • Teacher (Anna) • Manager (May) • Students 	provenance of the evaluation policy 6.4 Analysis of how the wider higher education sector frames the evaluation discourse.
RQ1.6 The experience of one particular student	<i>Classroom episodes</i> <ul style="list-style-type: none"> • Negotiation of essay title • Evaluation event <i>Interviews</i> <ul style="list-style-type: none"> • Teacher (Anna) • Student (Laure) 	7.2 Analysis of one student's construction of, and contribution to her learning experience 7.3 Analysis of this student's response to the different materials types introduced in the programme 7.4 Analysis of her experience of negotiation an essay title with the teacher.

Figure 4.11 Summary of research design for main study

4.11 Summary

This chapter has outlined the research strategy, as a plan derived from the experience of researching CS1, and the literature on qualitative research on the one hand, and as a reflected account of the process of researching CS2 on the other. The discussion relates the philosophical basis of naturalistic inquiry to the research tools of case study and ethnography. The concept of validity is discussed at two levels: first, the issues which I, as researcher considered and attended to, and second, the issues which relate to the policy and practice of programme evaluation by teachers. This latter point serves as a broad theme which encompasses the various strands of the programme examined in the analysis Chapters, 5, 6, and 7 below, and also as a guide for identifying ways forward for evaluation theory, policy and practice in Chapter 8.

Chapter 5

Programme constructs and evaluation processes

5.1 Introduction

This chapter presents the evaluation of CS2 - the main programme case study in this thesis - as a process of inquiry, in a specific educational and social context, and with specific purposes. The context, described in detail in 3.2 above, is a higher education English for Academic Purposes classroom. The purposes have been set out in two forms so far. First, a set of generic purposes deriving from the discourses of education evaluation examined in Chapter 2 above, and which can be summarised as accountability and development. Second, a set of grounded purposes which are generated by the pilot study in Chapter 3 above:

- a demonstration of quality assurance at programme implementation level (in many ways a higher order function of which the others are implementation strategies);
- a means of getting feedback from students on the appropriacy of activities for their learning needs;
- a means of getting students to reflect on their language skills development, and identify what they still need to learn;
- a means of persuading students to engage with the opportunities for learning outside the classroom;
- a means of negotiating emphases for the remainder of the programme.

The approach to analysis of the EAP programme evaluation in the context of the overall research design and specific research questions is set out in 4.10 above.

The analysis in this chapter focusses on the links between:

- the programme and the evaluation procedure;
- the evaluation procedure (design) and the evaluation event (what actually happens);

- the event and the consequences, both immediate for the programme in question, longer term for the teacher and future programmes.

The analysis of these facets of the programme in this chapter serves three overarching goals in this thesis. First, it determines evaluation validity (see 4.9 above). Second, it establishes a base for examining the macro-context (the higher education institution and wider sector) in Chapter 6, and a micro-context (the programme experience of a particular student) in Chapter 7. Third, it provides a platform for consideration of future development of evaluation theory, policy and practice in Chapter 8.

5.2 The programme - three perspectives

a) The specification of the institution

The programme evaluated here as CS2, is a formally approved one in a Higher Education institution. The procedures for approval here are those shaped by the Council for National Academic Awards (CNAA) in the 1970's and 1980's - a focus on close examination in committee of the construct of the programme by the university teachers involved and peers from other subject areas within the institution and from the same subject area in other institutions. The formal document on the programme is not easily available to students: instead the team 'translate' the key points into a more student-friendly handbook, called the module guide. The section on *Content* is included as Figure 5.1 to illustrate the construct which for the EAP team of teachers, represents the nature of the programme to both potential and actual students.

Module content

Each class will be composed of a variety of speaking, listening, reading and writing activities in Standard English, including:

- a) Discussion of articles selected from a variety of sources, including newspapers and journals, audio and video tapes. To gain practice in seminar skills, you will be encouraged to prepare a text distributed, and discuss it with the rest of the class paying particular attention to developing a critical analysis of the content of the text together with an understanding of its vocabulary, idioms and structures.
- b) Specific inputs on different essay styles, academic writing or study skills aimed at helping

you to manage the progress you expect to achieve in your main body of study. Relevant tasks in groups or pairs.

c) Opportunities for you to check corrections or comments made on your written homework, bring in assignments from other areas of study for general comprehension assistance or comment; work in groups to plan the written homework set for the following week.

Module Guide page 1

Figure 5.1 Construct of the programme as represented in the Module Guide

b) The construct of the teacher

When an approved programme is actually implemented there is inevitably a process of interpretation of this formal construct. This is evident in Anna's introduction to the programme as a pedagogic experience in the first session:

Anna then introduces the content of the module: what we are going to do. The general aim is to help students do what is required in a British university. Also English for you - as people who have spoken and lived with other languages.

In class we integrate the four skills - Listening, speaking, reading, writing - into activities. I am going to make you talk - you will be confident at the end of the module. [...]

Anna talks about the importance of speaking in British education: Students' programmes or other modules may involve them in speaking - you may be assessed in other modules on speaking.

Reading - Anna points out how reading may be hindered by cultural content. [...]

I will draw attention to the cultural content. As an example of what is involved here - metaphor. Often the literal reading is not enough. Anna tells students they will need to read a lot.

Academic writing also pointed out as very important - developing argument; critical analysis - key point in British degree programmes. Emphasises that writing is not descriptive - as in other places where student may have studied - they just have to reproduce facts. Not like that in Britain, opinion on the work of others is required. This may seem strange, but you need to pass judgements and critical comment. We will spend a lot of time on that.

Anna also emphasises talking to fellow students.

Classroom notes - CS2:1

Figure 5.2 Construct of the module as represented by the teacher

Anna, in this introduction to the programme, ventriloquises the functional orientation of the programme - the attention to the four language skills. In addition to this, she adds a distinct contextual element, as she situates the programme in terms of the broader institutional (Higher Education) and cultural (British) context

of the learning. She sets out what they have to learn, not so much in the technical, linguistic terms of the module guide, but in terms of socialisation processes involved in their participation in a British University. Students need to develop a critical faculty based on who they are, an ability to make a point carrying the authority of experience and identity. The view of learning represented here is in line with current views of EAP education, for example, Simpson (1996); Ivanic and Clark (1997). It also resonates with a significant theme in general education, a view of teaching as development of a critical perspective, an ongoing process of inquiry into learning, as represented in the work of Habermas (1968; 1987) and Young (1992):

Open to inquiry means being aware of the processes that produced the knowledge, having *some* practice in open-ended inquiry for themselves, and/or awareness of the ongoing inquiry - the contemporary discourse - and some degree of access to the discourse. That means, of course, access to a community.

Young 1992:13 (italics in original)

An example of this orientation of Anna's pedagogy is evident in her explanation of how students can take a critical perspective to what they find in texts:

11.52 *Anna relates all these points [students' experiences with computers] to the issue of being critical in writing; raising doubts about what is in writing like this. Students can do this by emphasising real-life problems - very good in essays.*

Classroom notes CS2:1

The extent to which this component of the programme is realised and its representation in the evaluation processes are discussed further in Section 5.4 below. The classroom observation data and teaching materials offer a third perspective on the programme – what is actually done in the classroom.

c) *The series of classroom activities.*

Since all sessions were observed and described and all materials collected and coded, it is possible to elaborate a content analysis of the programme in terms of the learning activities involved. The analysis of learning activities in this section

merges the concepts of materials and classroom activity, a convention common in the discussion of communicative language teaching (Legutke and Thomas 1991; McDonagh and Shaw 1994). A summary of these activities is attached as Appendix 5.

To analyse the pedagogic orientation of the materials a binary distinction between i) a focus on skills which might be considered meta-EAP: i.e. the knowledge and skills which might be considered as directly transferable to the demands of studying other subjects through English, and ii) direct language skills practice: a focus on language use and form more generally, where learning might be considered to assist the study of other subjects through English only indirectly, through improved English language skills. Within category i) a further binary distinction is made, adapting the concept of analytic and synthetic language use, first elaborated by Wilkins (1976) and adapted by Lawrence (1990) for curriculum evaluation purposes (see 2.6c above). Analytic materials are those which examine the EAP issues in a single complex discourse, e.g. 5/3/1 (5th March, first activity in session) - a research report on the criteria lecturers use to assess undergraduate essays. Synthetic materials are those which present discrete elements of EAP discourse, e.g. 12/2/4, a list of expressions which can be used for signposting in essays, etc. With regard to the second category - the materials for general English development - the classroom notes illustrate in detail the skills focus of each activity. What is represented in the coding is however, a simplification. For example, the cultural knowledge relevant to EAP students who wish or need to be socialised into a British higher education community, and the cross references between meta-EAP knowledge and skills which permeate the classroom discourse, are not represented. The discussion of activity 5/3/3 - Personal Language History - in 5.4 below illustrates how a critical perspective can make what is on the surface a speaking skills activity into a meta-EAP learning experience.

The representation of the programme emerging from this analysis provides an overview of what the programme involved. The time spent on these activities amounted to 73.6% of the total (i.e. 1590 minutes out of a possible total of 2160

minutes of the 36-hour programme). Figures 5.3, 5.4 and 5.5 show the distribution of the material types across the programme.

Materials Type 1		Materials Type 2		
Meta-EAP		Direct language skills practice		
Analytic	Synthetic	S (Speaking)	RSV (Reading Speaking Vocabulary)	W (Writing)
15%	21%	13%	47%	4%

Figure 5.3 Programme materials Weeks 1-12

Materials Type 1		Materials Type 2		
Meta-EAP		Direct language skills practice		
Analytic	Synthetic	S (Speaking)	RSV (Reading Speaking Vocabulary)	W (Writing)
16%	35%	9%	31%	9%

Figure 5.4 Programme materials Weeks 1-5 (pre-evaluation event)

Materials Type 1		Materials Type 2		
Meta-EAP		Direct language skills practice		
Analytic	Synthetic	S (Speaking)	RSV (Reading Speaking Vocabulary)	W (Writing)
13%	19%	16%	52%	0%

Figure 5.5 Programme materials Weeks 6-12 (post-evaluation event)

The proportions of time allocated to the different activities during the programme suggest that the student evaluation data would focus on meta-EAP activities with a synthetic orientation, and the Reading-Speaking-Vocabulary (RSV) activities based on texts. The accounts of the evaluation discussions in the following section bear this out, illustrating that though much of the discourse is framed as learner needs, with a strong future orientation (as described in the context on CS1, 3.5b above), the students are reflecting the experience of the programme to date.

d) *Constructs in conflict*

The discussion above is useful in two ways: first, it illustrates two dominant discourses of EAP learning evident in the institutional specification of the programme i) a functional discourse, emphasising the four skills, and ii) the teacher's construct, a critical discourse, representing the development of an academic voice. These discourses are evident in the meta-EAP materials, the former in those with a synthetic orientation, and the latter in those with an analytic orientation. There is a conflict between the discourses, evident in a divide between the teacher's construct (see 5.2b above) of the programme and her classroom practice (see 5.2c above).

The evaluation data (classroom discussion and written evaluations) and data from the interviews show that students are highly appreciative of the synthetic meta-EAP materials. The interviews with Anna show that she is aware of the popularity of these materials (a reason for including them in the programme) and also of their limitations:

Anna: [talking about handouts on EAP discourse markers] One of the bits I want to develop - at some point I am going to do it, I don't know if I am going to have time to do it this module, but where I am using those words, you know, connectives etc., what I want to do is have examples from academic texts where those things are used, to give them, I've done bits and pieces before, I would like to do it more systematically. So an example of *listing* from a text, from any academic text, to show how skilled academic writers in English might use it.

Interview with Anna CS2:1

Anna's intention here is revealing. She is aware of the limited value of these materials which are so popular with students. She can see how value might be added - by making the synthetic more analytic. The constraint in achieving this is time (it usually is - see 6.2b below). The heavy use of the synthetic-type materials may constitute a form of washback: Anna knows that they get a positive response from students, and that extensive use of them in the early weeks of the programme, is likely to give the students a feeling of well-being about their EAP learning, and thus, secure a positive response in the evaluation. This strand of

teacher thinking is revisited in 7.3 below, where in relation to Laure, these putative hypotheses of Anna's reasoning: i) that students like synthetic materials, and ii) that these do not benefit students as much as analytic materials, are further supported.

This analysis of three programme constructs helps determine the validity of the evaluation. The key question is the degree to which the evaluation findings, inferences and subsequent actions (see 4.9 above) relate to the different programme constructs and discourses. The next section (5.3) examines the evaluation event, with an emphasis on the process by which findings emerge, while 5.4, looking mainly at the interview data (gathered for the purposes of this research study, not as part of the programme evaluation) traces the life of the critical discourse within the programme.

5.3 The formal evaluation event¹

This section discusses the evaluation event in Week 5 of the programme, the mid-module evaluation based on the Nominal Group Technique (NGT - see 3.2 above). First, the event is described as a narrative, then a series of issues which influence the operation of the procedure as an evaluation event are examined.

a) *The evaluation*

The evaluation event took place in a manner similar to that of CS1 (Chapter 3 above), as identified through the classroom notes:

10.13

Anna introduces the module evaluation. There are three questions - on the OHP:

- 1. What has worked well for you?*
- 2. What has not worked so well for you?*
- 3. What would you like now for the rest of the module?*

Anna sets out the time frame - just seven sessions left. Asks students to answer each question, and also, to see the questions as a broad guide.

*Anna: But you should be able to put something under each question.
Unless you feel that nothing has worked, well ... (laughter all*

round). Take a couple of minutes to think about that, then I'll ask you to compare with other people.

Students chat and note - nearly all seem to be taking down questions from OHP.

10.15

Anna: Now I want you to work in groups, to reach an agreement on what to say under each on the headings.

Anna gives reason for this:

I want to prevent a, an outcome where one person says there is too much of something, and another person says there is too little. In that case I won't know what to do. After the group discussion we will have a report back. Then I will go away with something clear.

Anna then organised the 17 people into three groups - 2 of six and one of five. Anna says that she will leave the room so that they can say what they like (some laughter).

Classroom notes CS2:5

As Anna left the room I placed a small tape recorder on the table of one group.

(See 4.6c above for a discussion of this particular instance of data collection in the context of the research strategy.) The six students taking part in this discussion can be briefly profiled as follows:

Students/ Gender/ Nationality	Brief profile	Number of turns in discussion
Arnie (M) French	The joker in the group, but quite shy, not very positive about his English and learning more generally. Attends throughout, eight weeks out of twelve (8/12) (see 'Me' profile - Appendix 7)	26
Della (F) German	Keen, co-operative, aware, very competent. Attends throughout (10/12)	23
Helen (F) French	Quiet student, competent, not very social in group. Attends throughout (8/12)	9
Laure (F) French	Keen, co-operative, positive, gregarious, consensual. Attends throughout (11/12) (see written evaluation - Appendix 7)	40
Rata (F) Thai	Student who is struggling, finding everything difficult, very quiet and reserved. Attends Weeks 2-6 only	6
Sao (F) Thai	Student who is struggling, finding everything difficult, always eager to try to explain why. Attends Weeks 2-7 & 9 only (see written evaluation - Appendix 7)	14

Figure 5.6 Profile of the participants in the audio-recorded small-group discussion

These students discuss the three questions for twenty minutes. There are three occasions when a second conversation develops, and is not recorded, and numerous instances of 'off task' behaviour - social chat, or joking usually initiated by Arnie. Figure 5.7 summarises the data generated in the recorded small group discussion. For the most part the discussion proceeds according to the structure set out. The boundaries between the three phases are clearly marked: Della starts off with the invitational question "What do you like?" Sao initiates phase 2 - what hasn't worked well - at a point where Arnie has led the discussion off task, and Laure introduces the third phase - suggestions for the rest of the programme after a similar lull. A small number of points are made outside the nominated phase (there are included in Figure 5.7 within square brackets).

Student	Comments - What has worked well for you
Laure	Link words for essays
Della	Topics like ethnic conflict
	Group work
	Teacher's feedback
Sao	('I think it is the vocabulary I don't know' - difficulty for non-Europeans)
Della	Meaningful texts
Arnie	(Texts difficult to read)
Della	(No activity like reading comprehension)
Laure	(I want to speak about my text)
Rata	(More about the texts 'in the succession')
	(Difficult specialist subject text)
[Arnie]	[I like Anna's way of speaking]
[Arnie]	[I like the break]
Comments - What has not worked well for you	
Laure	Speaking in front of the class
Della	Structuring speaking
Helen	(More speaking in front of the class)
Arnie	Problem is speaking
Laure	(I like Anna's way of speaking)
Helen	Three hours is too long in one class
	Texts too long to read in class
[Sao]	['I think it is the vocabulary I don't know' - difficulty for non-Europeans]
[Arnie]	[No activity like reading comprehension]
Comments - What would you like now for the rest of the module	
Laure	More work in groups
Helen	More writing
	Speaking

Della	Writing - not in groups
Sao	Listening comprehension
Sao	Writing, e.g. writing summaries
Arnie	(I like the break)
[Della]	[Need more speaking in front of the group]

Key: () = points made relating to a question other than that being discussed in their chronological place in the discussion
 [] = points made relating to a question other than that being discussed in the appropriate category in the table

Figure 5.7 Content analysis of the recorded small group discussion

The next stage of the event was the plenary report back and discussion. This was managed by Anna, pen in hand at the overhead projector, taking points from each of the three groups in turn. Figure 5.8 is a summary of this discussion.

Phase	Point from recorded group	Student	Point from other groups
1		Ina	Writing
1	Talking; discussion	Arnie	
1		Ina	Interesting texts
1	Linking words	Laure	
1		Sue	Teacher's language; speech
2		Mat	Oral presentations
2		Mat	Vocabulary
2	More time for reading	Sao	
2		Mat	People coming in late
2	More oral presentations	Della	
2	Writing homework	Sao	
3		Ina	Grammatical exercises
3	Three hours too long	Arnie	
3	More oral presentations	Della	
3		Sue	Explanations of vocabulary in texts
3		Ina	Rewrite corrected essays
3		Sona	Shorter texts
3		Sue	Vocabulary focus for writing in class
3	Planning in class	Della	
3		Joe	Show examples of good homework

Figure 5.8 Content analysis of the plenary feedback session CS2:5

A key feature of this stage was Anna's silence - she simply noted - until Sao makes a point about more time needed to read texts in class:

Sao: More time to read texts. *Anna takes this up to try to understand what precisely is the problem, or rather the solution:*

Anna: We can talk about what it is you are saying here, do you want more reading outside class? Do you really want short texts? Because short texts are not what you have to read for your main study.

Classroom notes CS2:5

For the rest of the time, approximately 30 minutes, Anna receives points from students and seeks clarification or comments on them.

When the topic data (Figures 5.7 and 5.8) from the two stages of the evaluation event are compared to the pedagogic focus of the sessions in the programme up to that point (Figure 5.4), a thematic coherence is evident. The synthetic meta-EAP materials (lists of discourse connectors) where the feedback was positive, and RSV materials (long text-based materials) where the views were more problematic represent approximately two-thirds of the class time up to this point. This analysis questions the suggestion in CS1 (see 3.5 b) above), that students' evaluation are to be future-oriented, rather than related to the actual content of the programme to date. The future orientation of the CS1 students may be a discourse strategy, where the programme experience is the focus of the discussion, but comments are framed as suggestions for the future. This may be politeness strategy on the part of students, where the risk of offending the teacher through some direct comment is avoided. It may also be a function of the decision-making and action orientation of the event: students were, perhaps, responding to Anna's reminder that "we only have seven sessions left", and relating comments on their experience to that.

The content aspect of construct validity is confirmed in the correspondence between the analysis of the construct of the programme in 5.2 above, and the summary of the evaluation findings in Figure 5.8. The dimensions of validity which relate to the management of the event in the classroom are explored below. These are i) the commitment of the teacher to the evaluation event; ii) the status of the classroom event as a language learning activity; iii) the equality of status of all voices; and iv) the role of the teacher in managing the procedure. These factors provide a basis for describing the social and pedagogical nature of the event, and

identifying the factors which contribute to its evaluation credentials and those which might be limiting its potential.

b) Commitment of the teacher

The NGT evaluation event in Week 5 was carried out in accordance with the evaluation policy within the department responsible for this programme (see Appendix 6). Anna and I had several conversations about the evaluation procedure in the days preceding the event.

The week before I had asked Anna if this was the day when she would do the evaluation. She said that it was Week 5 in the module guide, so she would do it. There was a question about whether the students were sufficiently into the module to talk about it in an evaluation. She set out how she would fit the evaluation in with the other things planned - the main one was an activity using different texts for students from different subject areas - clearly it was going to be tight.

Later Anna says that she is not sure that all teachers are still following the policy of mid-module evaluations as strictly as last year when the department was facing an external audit.

Classroom notes (Headnotes on informal discussion) CS2:5

I noted in my research journal at the time, that Anna seemed less than enthusiastic about the evaluation. In the initial questionnaire study of the EAP teachers (see Appendix 2) Anna had commented critically on the evaluation policy - it was 'a bureaucratic one-way street' (these comments are quoted and discussed more fully in 6.4a below). I was concerned, therefore, that i) the evaluation event might be taking place simply because I was researching it, and ii) the evaluation data might be 'thin' if the classroom discussion was rushed. One answer to my first concern is that the event had a degree of fit with the programme as a whole and with Anna's pedagogy. Therefore, it is unlikely it only happened because I was in the classroom (see also 6.4 below). With regard to the issue of thin data, I suggested to Anna:

The day before the evaluation I feel that it would be useful for me to have a written, reflective piece from each student on what they expected from the module and what they were getting - complementing the group feedback that would emerge from the evaluation session. I put the idea to Anna, who

agrees in principle that it would be useful, and agrees to ask students to do it to facilitate me, ... to let them start in class, and finish outside.

Classroom notes (Headnotes on informal discussion) CS2:5

Ten students returned texts the following week, commenting as asked on their expectations and experience of the module and the aspects of English they felt they needed to concentrate on. Two examples - those of Sao and Laure - are attached as Appendix 7.

Anna's commitment to the evaluation policy of the institution is flexible - she is aware of the prevailing wider context, of the need to be seen to be implementing this policy, but is more interested in the benefits for her own classroom. Two points made by Anna in an interview following the evaluation event illustrate the pedagogical and managerial functions of evaluation. First, Anna sees it as useful for teaching - criticisms are opportunities to develop the programme, to understand and meet students' needs:

Anna: Usually they are both useful as a kind of mental tug to get a sense of how people are feeling, but also comes this awkward point on which people, sometimes, are not sure what people have to say, and it is sort of trying to manoeuvre it to make it positive, and for example in this evaluation, one of the things that stuck out for me was, people saying the texts I gave them were too long, and I'm still thinking about that one, ...

Interview with Anna CS2:7

Later, Anna points out the uses of the evaluation outcomes to the management of the wider departmental context:

- I: Have you discussed it with the other teachers of this module?
T: Not yet, in fact normally what happens is, in normal times, we write them up, and then, in the past I have written them up, and then I've passed them onto the module co-ordinator, who passed them onto the Subject Committee, that's normally what we do. I am just trying to think whether we have had formal sessions for this module, I don't think we have had a formal session, I think we have had informal conversations about what the outcome was, if there was any particular thing that we think other persons should know about, you know like, in this case I probably would have said to

May [module co-ordinator], that a lot of students were complaining that the texts were too long, and I might have asked her to check if this was a problem with other groups.

Interview with Anna CS2:7

A key issue in understanding the nature of Anna's commitment to the evaluation is the degree to which it conforms to her pedagogy, her teaching style. She managed the event as she would any activity, she engaged with students in the plenary discussion as she tended to in RSV activities - using probing questions to highlight issues and challenge students to clarify their positions or views.

Her hesitations about evaluation seem to centre on use of findings outside her classroom; she is less than enthusiastic about reporting findings to the co-ordinator of EAP programmes or more remote managers in the institution. Her initial comments about the difficulty of fitting the evaluation in, and the fact that other teachers were not doing evaluations may have been intended to lower my expectations regarding outcomes, such as a report to managers. This view is consistent with an important contemporary discourse within education at all levels in Britain: resistance to bureaucratic or managerial procedures which absorb time and energy, but actually have very little value added benefit for teachers or students. It is an issue discussed in greater detail in 6.4 below.

c) *Language skills development*

Anna introduced the evaluation event as a context of language practice:

9.25 Then Anna returns to the evaluation - sets out two reasons/purposes: To give students an opportunity to say what they feel about the teaching and activities on the module, and secondly:

Anna: So as not to waste the time that we have, an opportunity to practise the language skills - you will make notes, discuss in groups, present points to the whole group, and finally do some writing.

Classroom notes CS2:5

The issue here is whether the harnessing of the procedure in this way lessens or degrades the primary, formal evaluation function. The view of the evaluation procedure as language practice represents a teachers' construct (not just Anna - the

discussion procedure was adopted as departmental policy because it facilitated language practice in a way that, say, questionnaires would not (see 3.2 above), the other EAP teachers surveyed at the outset of this study made the same point, and the report drafted by Millie at the end of CS1 (see 3.5c above) noted it). This function of the evaluation, however, seems one of rationale rather than experience - in interviews with Anna and five students in the period immediately after the event, there is no mention of its merits in relation to the language development function. One reason for this may be the fact that the language development skill involved in the evaluation event is speaking: although students often note that they need to improve their speaking skills, they tended, in interviews, not to note improvement in this area as a benefit of a given activity. The issue of transparency in learning needs and classroom solutions is discussed in detail in the Laure case study in Chapter 7 below.

The language practice issue may reflect a different discourse for the teachers. It may be that, in a curious way, teachers are stating the pedagogic function of the activity, to signal their reservations about the management or monitoring function. Thus, they establish a distance from a process which is regarded as bureaucratic and of little value, but without positioning themselves in direct opposition to it. Thus, in public pronouncements concerning evaluation, teachers may be stressing the pedagogic function as a means of signalling ownership of the activity. There may be a correspondence here with the issue of commitment discussed above: the teacher's instinctive response in justifying components of a programme is to relate them to students' learning. This is a core responsibility of the teacher, and one which students have an interest in and an awareness of. The extent to which this pedagogic perspective on the evaluation takes from or compromises its quality management functions is examined further in 6.3 and 6.4 below.

d) Equal status of all voices

A central feature of NGT as an evaluation procedure is that it works with the issues raised by students. The method of inquiry is essentially *emic*, rather than *etic*: the criteria for making judgements and decisions derive directly from students' concerns. This section explores the issue of equality of status for all

student voices through the analysis of a segment of data from the recorded small group discussion (see 5.3a above).

When the plenary discussion stage of the NGT starts, there is a problem in categorising points. There is an overlap between *what has not worked well* and *suggestions for the future*, a phenomenon observed in CS1 (see 3.5b above). Students tend to note *what has not worked well* as personal shortcomings, or what they do not like, but actually need. Della, picking up and developing a point raised by Laure, sets out this perspective quite clearly - the issue of practising speaking skills. The data illustrate a number of features of the discussion:

Della: ... I wonder if something could be done, so we have more speaking in front of the group, so that we do that, (laughter) I don't love it really, but ...

Laure: ... More speaking but in groups not in front of the class

Della: ... no, but in front of the class, I get a bit excited sometimes, it is quite useful, we can say. I think it's good.

Rata: Perhaps if you know this and are good at this ...

Della: ... No, no, I didn't like it before, because maybe like you, I didn't know the structure, but now, it's a good exercise, and if it is not good nobody cares, we are all the same, and nobody comments, I think it's fine

Laure: Yes, and in fact it will be good with the (Unclear) oral presentation

Transcript of small group discussion stage of NGT CS2:5

Della, a strong student, proposes more speaking, saying that she needs it but finds it difficult. Laure, the consensus seeker (see Figure 5.6 above), reinterprets Della's preference as a desire for more groupwork, which Della quickly clarifies. Rata points out - or tries to point out - that the situation may be different for weaker students. Della, noting that she is being considered as a strong student, insists on identifying with the others. Then she claims, revealingly, that different levels of performance do not matter. Laure takes the opportunity at this stage to make a general point - relevance to an assessment task - but not a helpful one in understanding what has not worked well, from the perspective of Rata.

The point made in the discussion of the evaluation procedure of CS1 - the lack of analytical focus on the precise nature of problems - can be observed here. There is however, some evidence of an explanation: the social nature of the discussion. Those students who are strong, and benefiting from the activities so far, for example Laure (but see Chapter 7 below) and Della, are keen to represent themselves as having personal shortcomings. This may be due to natural modesty: they do not wish to display what might be interpreted as arrogance or superiority, but rather to establish common cause with the weaker students. This pattern of behaviour has two outcomes - it avoids exploring differential effects of classroom activities. Perhaps the stronger students have a sense that such a discussion would be divisive, and negative in terms of the task in hand. Second, in the search for consensus, it is the concerns of the stronger students that prevail in the group, neutralising the impact of the serious points raised by, for example, Rata. This gives a stronger voice to the students satisfied with, benefiting from, and keen to maintain the status quo. Making suggestions is a discourse function which facilitates this better than scrutiny of negative experiences of the programme to date.

This strong conclusion from an analysis of a small piece of data is supported by other evidence. First, the points on which the students are negative or critical relate to issues which are procedural, with little bearing on the substance of the programme. These points are made by strong students - competent participants in the classroom discourse. Second, the experiences of the Asian students in the group - Rata, Sao and Niki - seem qualitatively different from the other students, and more importantly, absent from the consideration of points in the evaluation event.

The negative experience related by Sao has a cultural dimension: she feels that part of the problem is that she is from an Asian culture, and is thus disadvantaged by the European orientation of the programme.

Sao: I think this module is too difficult for me (laughs) because, if you come from a country like Thailand, outside Europe, you can't,

communication is difficult, but if you hear something, because I know, I just know about how the course of the module is emphasising how to read the words, how the combination, no, and all the other things

Transcript of small group discussion stage of NGT CS2:5

Her problem is instantly ‘solved’ and thus eliminated by Laure, who represents the classroom as a community, without realising that this is precisely what is not working for Sao:

Laure: But you know as most of the time we are working in groups, if sometimes you don’t understand an expression or a word, you can ask the other people in the group to explain you the word or to explain you the general idea, for that it is worth it, we are in groups so we can exchange our point of view and help each other (Sao indicates agreement as Laure speaks)

Transcript of small group discussion stage of NGT CS2:5

Sao’s view is echoed by Niki, a Japanese student on the programme. Her strong feelings do not emerge in the context of the evaluation discussion, but in an interview, and relate not only to the programme in question, but to a wider context of English language programmes that she has experienced². Niki makes it clear what these feelings are: she explains why she does not like groupwork and collaborative activities in the classroom:

Niki: Because it hurts my feelings, if I mixed with the European peoples’ group, the European people always look down at the Asian people, because they believe that Asian people cannot speak English well, or cannot understand English properly, so it’s a bit like racism, this idea, but I sometimes feel a bit uncomfortable with questions, so I really hate that, so I don’t like, it really depends on that, what kind of group and what kind of people are involved in the group, but I sometimes feel so uncomfortable with doing that groupwork.

Interview with Niki CS2:7

The situation set out by Niki above is well known in EFL teacher folklore as a classroom problem, but one on which the ELT research literature is somewhat silent. Senior (1997a) examines the group dynamics of adult EFL classes with

Asian and European students, but discusses problems with role and social dimensions from the perspective of the teacher only. She recommends that 'teachers should, a few weeks into an course, make a special effort to draw into the class group those students who still appear to be on the periphery' (1997:9), without exploring what on the periphery involves, why students might be located there, and the extent to which drawing students into the group is actually possible. These have not been looked at systematically within Applied Linguistics - a point made in Breen (1985); Bailey (1996); Woods (1996) and Senior (1997b)³.

It is evident here that a set of problematic experiences do not emerge in the evaluation. There are two possible reasons for this which relate to student interaction and reasoning. First, the strong students with a positive view of the programme dominate the discussion, such that it is discursively inappropriate (as well as linguistically and culturally challenging) for less able students to have their say. Second, the weaker students, where there is a strong cultural divide (e.g. Asian and European) and an accompanying sense that the programme is not benefiting them because of this cultural identity, collude in their own marginalisation. Rata and Sao try in vain to make the programme address their needs, but Niki, a more experienced student has given up; she has a developed view of the deep roots of the Asian-European divide in classrooms, and judges it best not to disturb them. She stays with the programme to the end, but Rata and Sao drop out after Week 9. They also drop out of their main programme of study at the university.

An interesting aspect of this particular problem is the extent to which it is managed by the students. It looks as if the strong students have conspired to shape feedback so that the level of activities continues to suit them. There is no suggestion of open collusion here: rather shared interests have, separately and silently, generated similar responses. Equally interesting is the teacher's role in the shaping this aspect of the evaluation discourse. One might hypothesise that Anna has short-circuited the emergence of problems which affect minorities in the group by structuring the discussion around agreed points. Anna was aware of the problems these students had - in the first week of the programme she told me that:

... many different needs seem apparent. There are some very articulate European students. Anna is aware of a potential problem here: students like this can appear confident, or arrogant and the result is that other students seem intimidated, and end up not saying very much.

Classroom notes (Headnotes on informal discussion) CS2: 1

Anna's perception of the potential problem here is in the behaviour of the strong students, and there is evidence in the classroom observation data that she takes a directive role to make sure that they do not dominate. This is discussed in 5.3e below. She also notes the problem of very weak students, and sees this as a problem with institutional policy - recruiting students whose level of proficiency in English is too low for them to benefit from her classes, especially the current, higher level EAP class. Sao and Rata seem to be the students who fit this bill, although Anna did not name them. She does however challenge a point Sao makes (the second time this happens - see 5.3e, page 134 below):

Sao: Writing the homework, for me it is very difficult.

Anna: Is that something from the group as a whole, is it something general that I can act on?

Sao goes on to expand on the difficulty that she has in writing, in expressing her understanding of ideas.

Anna: Is that a view from the group?

After a short silence Della takes it up, and mediates:

Della: Writing is difficult, but it is something that we need to practise, it is not easy, but it is good for us.

Classroom notes CS2:5

"Something general that I can act on" may be a code for an input factor - Sao's low level of English language proficiency - rather than a programme implementation issue. The resolution comes by way of Della skilfully acknowledging Sao's problem, and reconfiguring it in a way more acceptable to Anna.

Classrooms are 'status-marked' settings (Edwards and Westgate 1987:42), with an obvious and well-researched divide generated by the role of teacher and student. McHoul (1978) distinguishes between talk which is organised as it goes along, where 'none of the participants has any special rights or obligations to take the

decisions necessary to achieve orderly interaction', and talk which is organised 'largely through unequal rights and obligations arising from status differences between the participants' (1978:185). Much of this discussion in the education literature focusses on universal dimensions of classroom discourse - the adult-child element of roles, and the numbers issue, the fact that there is one teacher and several students. While these are not obviously salient issues in contexts with more mature students the issue of status is a real one. It not only profiles the role of the teacher, but divides students as well. Roberts, Garnett, Kapoor and Sarangi (1992) in an ethnographic study of further education classrooms conclude:

The extent to which students feel and act as if they are liked and respected, and the extent to which students appear likeable to teachers, clearly affects their success. [...] Teachers early on label students more or less positively and this labelling depends on social relationships in the classroom [...] negatively labelled students find themselves with fewer classroom learning opportunities.

Roberts, Garnett, Kapoor and Sarangi (1992:64)

One tentative conclusion to this discussion is that Anna is aware of a Pandora's box of what are perceived as unresolvable problems. Her response is to assume equality of status - evident in her overall pedagogic approach - while aware that it may not translate into equality of voice in the evaluation discussion. The data suggest that there may be other processes within the programme for dealing with issues which relate to individual students, such as responding to written work, seeing individuals during the break, and providing guidance for the assessed tasks towards the end of the programme. This provision by Anna of a kind of safety net is explored further in the context of student perspectives in Chapter 7.

The analysis of unequal voice status relates to the validity of the evaluation in two ways. First, in terms of the construct, there is a problem with how Anna operates the NGT (see Figure 6.2 below for an analysis of how Anna's procedure differs from the departmental policy). Important points relating to experiences within the programme are not discussed. These problems may be to do with components of the programme - the cultural identity of some participants, and associated prejudices - which the evaluation could not be expected to resolve, but which it

might be expected to engage with. Second, there is a problem with consequences (see 4.9 above): two students, arguably those in greatest need of improving their EAP, drop out of this programme and their university study before the end of the semester, in all likelihood due to an inability to cope with EAP demands. This latter point goes to the heart of what the institution wants from the evaluation of programmes and the quality assurance system as a whole: the success of students as represented in progression and completion rates. The quality issue here is deeply embedded in the concepts McHoul (1978) identifies as significant - rights and obligations. These determine interactions and accountabilities across the institutional community, particularly in the teacher/manager interface. The issues here are discussed further in Chapter 6 below, and in Chapter 8, where the issue of marginalisation of Asian students is explored to illustrate the potential of a multi-faceted, institution-wide, inquiry-based approach to quality management..

e) The role of the teacher

Anna's style of teaching is strong on leadership. She manages the classroom in a smooth, authoritative way, giving structure and cohesion to the long three-hour sessions. She has a clear view of what learning in the programme should involve, and is direct in communicating this to students (see 5.2b above). Her leadership style is complemented by a natural, teacherly bearing in the classroom which students respond positively to. In the evaluation, Arnie compares Anna positively with another teacher:

Arnie: I like her way of speaking, because you know [*names another teacher*], he speaks very, very low, never with expression. Here the teacher speaks very loud, she speaks with her hands, she is very, she is living [*lively*].

Transcript of small group discussion stage of NGT CS2:5

In the plenary discussion, Sue comments

Sue: I think your language, your speech; we can understand you quite well. *Anna laughs lightly at this, and there is a murmur of confirmation from the other groups.*

Classroom notes CS2:5

The NGT evaluation procedure, by contrast, requires Anna to play a different role, a more facilitative one, listening rather than directing. The question for this section is how these potentially conflicting roles are resolved, and the degree to which the resolution contributes to the validity of the evaluation event.

The issue here is power, the shared awareness of the status of individuals or groups which underpins their 'ability to achieve aims or further the interests they hold ... at the expense of others' (Giddens 1989:729 cited in Murphy and Rea-Dickins 1998). Considered in the light of the above discussion on the concepts of *status* and *authority* within the programme, we can identify the *power* issue as pertaining to some dysfunctional or undesirable outcome of the social structures within the programme. Given the divide described above within the student group, the question is the extent to which Anna legitimises this and thus allows it to shape the evaluation event and the points for action which emerge.

The evaluation in CS1, described in 3.5 above, illustrates the toggling in the role of the teacher between *listening*, as might be expected from the nature of NGT, and using the procedure as a *bully pulpit* for the teacher to set out her views, and persuade the students towards some desired course of action. In the evaluation of CS2, these roles are evident. Anna clearly indicates that she is not to be part of the initial stages - and she leaves the room for the period of the small group discussions. Anna returns to manage the feedback stage, and there is a clear pattern in her role here. During the discussion of Phase 1, she accepts the points of the students in silence, sometimes repeating them, then writing them up on the OHP. Even though the students' feedback does not proceed smoothly - not all groups have organised their points to facilitate feedback - there is an abiding element of ritual about the procedure, as if the roles being played have been practised and honed over time.

At the start of phase 2 - *what has not worked well for you* - Anna's role is less silent. The issue of text length is an example:

Sao: More time to read texts.

Anna takes this up to try to understand what precisely is the problem, or rather the solution:

Anna: We can talk about what it is you are saying here, do you want more reading outside class? Do you really want short texts? Because short texts are not what you have to read for you main study.

Sue: Maybe not so long. Maybe 4 pages, not 5.

Joe: Shorter texts when we have to talk about it in the class.

Ina: Yes, for homework longer texts are OK.

Anna: OK.

Classroom notes CS2:5

There appears to be a structure to Anna's response: seek further information from the student, require that this information be about the solution rather than the problem, note the points about this solution as they come. From the perspective of Sao, this exchange must seem like Laure's response to her problem raised in the small group discussion in 5.3c above: the problem which is marginalising her, is taken over by others (who are not marginalised), and leads to solutions which only serve to leave her on the periphery. Anna's probing of issues such as this, her deliberate listening posture, seems to create a vacuum in terms of ways forward, which stronger students fill with suggestions. Anna's 'O.K.' represents a 'shelving' of the issue, an acknowledgement that the proposed solutions are perhaps unsatisfactory, but are all that are available. Anna comments in interview later:

Anna: ... in this evaluation, one of the things that stuck out for me was, people saying the texts I gave them were too long, and I'm still thinking about that one, because the teaching sessions are so limited, twelve sessions or less really, if you take into account the oral presentation session, to give people just very short texts, with, small activities, doesn't seem to me to be matching the level of performance they have to secure in their main programme, and given that this is an assessed module as well, it seems to me that, I feel in my bones that I should be pushing them a bit more, than just simply giving them a series of short texts, with relatively unproblematic tasks, and for example if you are studying language, then, you should be able to read a four or five page thing about language, and take something out of it, and make some response to it, if you can't do that, then it will raise further problems about, well how do they handle that programme, I don't know, and also, because of the limited teaching time, I always feel I need to give them something, that they could, those who wanted to, could do outside class, they could go into it more, they could take their time,

you know, I could start them off in class with a long text, and pick up some points, and do some activities, and initiate some things in class, but I feel they should then take it away and do something more with it, and normally I'm not asking them to complete anything other than reading, a text, and normally, sometimes I'm even worried that I may be giving them too long, to read them, but essentially I give them at least 20 minutes, or even more, to read, and get the gist of a text, and then if they want to they can take it away.

Interview with Anna CS2:7

Taken together with the classroom data, it is evident that for Anna the issue here is the nature of text-based activities within the programme. Her rationale for long texts is complex and coherent, encompassing broader issues than the surface aspects of classroom activities raised by students. One might hypothesise about the reasons for her somewhat abrupt closure and shelving of the issue: a feeling that presentation of her rationale for long texts would not succeed; a sense of a firm position on the part of students; an express effort at playing a listening rather than persuading role; or a fleeting notion that her thinking may need revision (see 5.5 below for an exploration of a related text-based issue emerging from the evaluation: the question of how best to deal with new vocabulary in texts).

What is evident in Anna's management of the exchanges in the evaluation event, and in her subsequent reflections, is an assumption of homogeneity among the student group. This seems to be reinforced by the 'equality' discourse generated by the strong students, for example, the exchanges of Laure and Della in the small group evaluation (5.3d page 126 above). The effect for weaker students is to construct and sustain their invisibility in the programme. Time, coming in as a discourse of limitations rather than one of opportunity, plays a recurring role here. Anna's awareness of these problems seem not to suggest a solution, and leave her with no option but teach those who 'fit' the programme, and effectively ignore those who, in an ideal world, would not be on the programme at all. This is another classroom problem which is well-known in EFL teacher folklore. In the literature, it is recognised as a strategy for teaching large, heterogeneous classes, e.g. Shamin (1996). In higher education, policy and teaching developments have

through the 1990s focused on the importance of meeting all students' learning needs. There is some evidence that English language and related academic skills are a significant part of the problem here (Cortazzi and Jin 1996; Simpson 1996) and that it is a problem with a recruitment policy dimension (Furedi 1999; McKinley 1999; Doloughan 1999; Reyal 1999). There is, however, little research on the programme processes through which such problems are resolved, or perhaps more importantly, are not. Chapter 8 below considers how evaluation data available to Anna might be better used to inform institutional processes.

The role of the teacher in the evaluation is constrained by a number of factors: the uncertain space between an authoritative, directive style and a listening, facilitator approach; the void between students' simplistic characterisation of issues and the teacher's awareness of their underlying complexities; the difference between the marketing of the programme, and what is actually achievable in the classroom.

The factors seem to converge to generate judgements, the representation as positive or negative, with little description of or accounting for the actual phenomenon or experience being judged.

5.4 The values underpinning the programme

There is, as presented in 5.3 above, some evidence of a veering away from significant issues in the NGT procedure, a co-constructed avoidance of open discussion of factors which are impeding the learning of some students. This avoidance, it is suggested, amounts to a serious threat to the validity of the evaluation. This threat to validity may be compounded if avoidance of a critical perspective also compromises a focus in the evaluation on Anna's construct of the programme (see 5.2b above). This section, therefore, examines how the key values which inform this construct are evident in the implementation of the programme, and how they are evidenced in the evaluation itself. A related dimension of consequential validity (Messick 1989; 1996 - see 4.9 above) is the extent to which the awareness-raising function of the evaluation, and more widely, of the programme as a whole, is achieved.

This aspect of validity is explored through examination of one particular activity in the programme which represents a link between the classroom context and the wider social context, i.e. socialisation into the discourse conventions of British higher education. The activity - Personal Language History (5/3/3), done in class in Week 4, the session before the NGT evaluation, is a direct language skills task focussing on speaking skills, with a meta-EAP dimension (see 5.2c above and Appendix 5). Anna outlines the pedagogic rationale as follows:

Anna: Well, [...] the Personal Language History work intended to practise skills like questioning, oral skills like questioning, discussing, explaining, it's also of course a chance to practise taking notes of key points very economically, and it's an opportunity for them to have a different kind of delivery, in other words they are talking in pairs, interviewing each other, then the next stage is where they give an oral report back, or selective people do, in plenary, again oral work, and then ends with a written task, and in terms of how people performed in it, I don't think it was very variable, and on the whole, people are relatively unquestioning about language use, and in a sense see it as unproblematic a lot of the time, but in comments that I made in their written work, I attempt to put questioning in the margins, and to indicate where they've been very flat in what they said, you know, because again, that's one of the skills we are trying to practise in the module, to try to say that, well, you're in an academic arena, you need to go beyond a flat description of phenomena, and to be able to find the difficult areas in the problem, so the only place that I think I can do that in these exercises is in the margin notes that I make on the written assignments, because it's a much longer process to get people to understand why you think there's something complicated about the language use. It's clearly something that people don't find very obvious. [...] And what is interesting to me, teaching this module, is quite a few of the people, well everybody is at least, bilingual, and bilingual to a high degree, in this sense, I think everybody in the class comes from a country where they learn to read and write in another language, have got a high level of education in another language, and then transfer to English, and have achieved a high level of spoken, understanding of English, and that English, such that they can survive at undergraduate level, so they are all at least bilingual, and somehow or other multilingual, and also several of them are people actually studying language, and I find it a continually useful exercise, because given that these people have such a high degree of interest and expertise, they still don't see much complexity in the whole situation, and, [...], I find that quite interesting.

The question explored here is the extent to which Anna's rationale is evident to students, and informs their evaluative comments. With its surface-level skills focus and underlying meanings of learning as socialisation, this activity serves as a kind of metaphor for the programme as a whole. Anna's words: 'in an academic arena, you need to go beyond a flat description of phenomena, and to be able to find the difficult areas in the problem' resonate the key values of her programme construct discussed in 5.2b above.

In the evaluation feedback, the small group discussion and the plenary session, there are no references which might be identified as relating to the development of a critical perspective and skills. I asked a number of students to comment on the activity in interviews in the weeks following. Students' comments reveal four kinds of response:

- i) no sense of the value of the activity

Isa: ... maybe she wants to know our background, our language background, [...] I think she has a lot of attention in some points that maybe is not very important, or, I don't know, I think sometimes that she put the attention on some few details, for example, the dialect ...

Interview with Isa and Jamie CS2:4

Niki: I just think this one was a personal language history, I think it is just for, I didn't relate it to my academic skills.

Interview with Niki CS2:7

- ii) a sense of value for language skills development only

Jamie: I think in my opinion, this lesson is useful practice, practise and practise, and develop speaking, and of course, language points I don't know but when I arrived, I was expecting to have some English grammar, more basics, English basics

Interview with Isa and Jamie CS:4

Sara: Develop tense forms, make us talk to each other, and to note down the things ...

Interview with Sue and Sara CS2:5

iii) a sense of value for language skills and self awareness in relation to language

Ina: Oh I love it, doing it, because, when you have to write an activity about yourself, for me that is the easiest thing, because I didn't have to look information or nothing for it, I have to think about myself,

Interview with Ina CS2:6

Sue: This was really interesting. My problem is that I am only speaking German. My mother isn't German, but I don't speak my mother's language at all, and then I heard all these other people, how they are able to speak that language, and that language and that dialect, maybe ten languages. But it is incredible, and so interesting for me, such a wide range of languages. I really liked this activity, it also asked for, things like, not only the obvious things, but things like slips of the tongue, like attitudes.

Interview with Sue and Sara CS2:5

Della: ... but my personal language history, for example, was useful in another context, because I thought about how I learn languages, and I thought about how could I do something better, maybe more effective. And this was very useful, yes.

Interview with Della CS2:7

iv) a sense of value for development of a problematising, critical faculty

Joe: Topics about language, I think were very useful, because when I tried [to] write something about, it's not the essay, I try write the essay, and I find quite difficult and afterwards when teacher came and told me how to do it, and how should I look inside, how many different languages can be, it is not only that it is one language and what is the culture that the language is changing, and in our country, it's not so much problem, because only one spoken language, and, one official language, but some regions they have different, like dialect, they are talking, they use different words, and sometimes they are difficult to understand, and I was thinking

about language, and particularly in my country, I thought it was only one language, when I didn't take any different aspects in that region [that] can be different ...

Interview with Joe CS2:7

Annie: Yes, it was useful, because that writing is very useful, because it helps with things you take for granted, it's really useful for that [...] Like, how other people, how they communicate, some are very quiet, not very much talking, and I could understand how different people speak, how people use language in my country, that was interesting.

Interview with Annie CS2:9

This range of views, where the same activity in the classroom is seen as one involving bureaucratic form-filling where they provide information about themselves, through conventional language skills practice and activities of personal or topical interest, to one which leads students to a new perspective on their own experience and related social phenomena. These views can be informed by a theoretical framework for critical learning, based on a pedagogic reading of Habermas' Theory of Communicative Action (Young 1992:39-40)

Young sets out four kinds of communicative action, which represent different kinds of interaction - a process of conceptualising and responding - between the self and the other.

i) Strategic action, where there are different interest positions, and these are maintained, agreeing to differ, and continuing to achieve previously established ends. Young describes this form of action as cognitive-instrumental, with critique limited to noticing a failure of ends/means matching. In the case of the evaluation explored here, the view that the programme is about providing unimportant personal details only, in spite of different messages from the teacher and activities in the classroom, represented to some extent by the view of Isa.

ii) Normatively regulated action, where there are similar interest positions, and thus a common frame of reference, encompassing 'common values and norms, instilled through socialisation'. The form of critique found here is based on conformity and non-conformity: in this context this approach is evident in the ventriloquising of EFL concepts and norms, e.g. the points raised by Jamie and Sara above. Anna's view of the programme is to familiarise students with other norms, as in iv) below, in order to allow them to be competent participants in higher education, not just in the EFL classroom.

iii) Dramaturgical, or expressive actions represent an initial personal response to the other, possibly the recognition of a mismatch between the learning activities and personal feelings. The student is in the role of actor, experimenting with other voices, tentatively articulating ideas which are new, and both interesting and confusing. The responses of Ina, Sue and Della above illustrate in different ways this form of action: the personal expression is very strong, but there is as yet little certainty about how this questioning voice might relate to other domains.

iv) Communicative action represents for Young, as for Habermas, a pinnacle of learning. It is a 'culturally encapsulated critique', i.e. dealing with all aspects of the context in question, 'relativised against the possibility that validity will be contested by others' (Young 1992:39). Each participant holds a view which has itself been critiqued, but is still open to new critiques. Equally the critique of the other view is informed by the context, the interests of the other as well as the interests of the self, and the possibility that it may be valid in ways not yet considered. Joe and Annie represent a questioning of their social context and their own experience within it, which illustrates achievements in critical learning. They also reflect positions on the social realities they refer to which relate to the fundamental goals of the programme as interpreted by Anna (see 5.2b above).

The critical faculty, a central element in learning in higher education, has many facets and representations. Anna emphasises it as something 'British', the desired alternative to repeating the lecturer's notes or readings from the literature. The paper she uses as materials in the classroom (materials 5/3/1 and 16/4/2, based on Riddle (1994)) relate 'being critical' to 'making argument' and posits 'critical analysis' as a strength in student essays, where 'mere description' is a weakness. Lea and Street (1998) in a study of writing requirements in higher education which explored lecturer and student interpretations report that the former places a high value on 'being critical'. The study found that while students are aware that 'critical analysis' and 'coherent argument' are important, they struggle to understand what they actually mean. Lea and Street's conclusion is that the notion of critical faculty, as liberally used by lecturers, remains an undescribed and unexplicated concept which is embedded in constructions of knowledge within subject disciplines (1998:163). This perspective helps understand how being 'critical' is a problem for an EAP classroom such as CS2.

In CS2 Anna has students from humanities, social sciences, health sciences, law, accounting and information technology. Her critical pedagogy, based on their experience, and relevant to their needs has to bridge two divides: i) writing in English generally, and writing for their subject area; and ii) academic writing in their 'home' context, and in Britain. Her achievements as represented in the interview data are not insignificant. The question here is why this critical aspect of the programme does not feature in the evaluation. This phenomenon may constitute a threat to the validity of the NGT procedure, in much the same way as the issues discussed in 5.3 above. The more fundamental issue is the capacity of students as evaluation informants, where they have only a partial and vague understanding of the complex goals and sophisticated designs of programmes. This is taken up again in Chapter 6, as a problem where institutions rely heavily on student feedback in evaluations; in Chapter 7 where one student's programme construct is examined in detail; and in Chapter 8 where recommendations for improved practice are considered.

5.5 Some consequences of the evaluation

The preceding section explored the concept of validity of the evaluation from the perspective of what might have come up, on the basis of an analysis of key components of the programme. This section starts with what did emerge as evaluation findings, and trails one issue, vocabulary, through the remainder of the programme, in order to determine the actual impact of the evaluation. Three separate types of consequence are relevant here:

- i) what happened in relation to stakeholders beyond the classroom;
- ii) what happened in the classroom; and
- iii) what happened in terms of Anna's thinking about the programme.

The data inform on the two expected outcomes of the evaluation. In terms of accountability, the reporting to colleagues or the EAP programmes co-ordinator is vague: Anna is aware that this is something which should be done, but by the end of the programme no report has been produced. There have been informal conversations within the team of teachers teaching different EAP groups, though these may be more concerned with administrative issues- clarifying class lists and procedures and schedules for assessment - than pedagogic ones. There was no summative evaluation, a breach of departmental policy (see Appendix 6). The problems with the reporting aspects of evaluation policy are taken up in 6.4 below.

The developmental outcome, or consequence (Messick 1989; Stake 1995) of the evaluation is evident in the data in a number of forms: the classroom data show Anna using it as a reference point in subsequent sessions, for example a reference to the problem of sufficient time to read long texts :

9.45 Anna announces the next activity - a text to work on:

Anna: ... this time I'm taking into account what you said last week, I'm giving you plenty of time to work on it.

Classroom notes CS2:6

Eight students referred in detail to points which arose in the evaluation in interviews with me after the procedure. Most significantly, Anna was able, in interviews, to refer to specific details of the discussion, to illustrate how this feedback related to feedback from other groups in the past, and to discuss the issues in relation to what she was trying to achieve in the classroom (see, for example 5.3e page 134-5 above). To explore in detail the consequences of the evaluation in terms of programme development, the issue of ‘vocabulary’ is selected for three reasons. First, it is a term that students use when talking about learning and teaching, socialised as they are in the discourses of EFL. Second, vocabulary seems to serve as a code for the specific problem of plagiarism: when students talk about expressing an idea or a view in their own words, they mean how to avoid simply copying what is in a published text that makes the point very well (see the discussion of Abe’s problem in 3.5b above). Third, there is a vagueness about the term when it is used by students to describe issues in foreign language learning. It is a problem that will always be there for the language learner; a source of frustration particularly for advanced learners who need to express complex concepts and ideas; a problem for which there are few learning strategies, and fewer teaching strategies. As an EAP teacher, Anna is aware that all students will state a need for more help with vocabulary, although the particular words needed or causing difficulty will vary from individual to individual. She has, thus, to mediate and interpret what is afoot when the term is used, and develop an appropriate pedagogic response. Thus, the notion of vocabulary provides a window through which to examine the interwoven processes of interpretation and action as they relate to evaluation and pedagogy.

a) Characterisations of vocabulary

‘Vocabulary’ was a point raised by several students in the evaluation in Week 5. Sue, a competent student makes a strong case for more attention to the vocabulary of texts in the evaluation plenary discussion:

Sue: More explanations about the vocabulary in texts
She gives a detailed account of the difficulty she has in dealing with a lot of new vocabulary.

Anna: Any suggestions? I don't want to make the session unbalanced, spending a lot of time going through vocabulary, but we could have ten minutes in each session where you bring up words, vocabulary items such as phrases, etc.

Students nod at this suggestion.

Anna: There are other ideas, such as keeping a notebook with vocabulary. I have suggested this in the past but the students have not taken it up.

Sue: We need to look at the vocab from the text, because it is so difficult to find out all the new words.

Anna: And when we have different texts? Any more ideas?

Classroom notes CS2:5

Sao, a student who seems challenged by the demands of the programme generally makes a similar point about vocabulary in the context of listening comprehension (see excerpt from small group discussion in 5.3d page 127-8 above).

In both instances the students are using vocabulary to label what might equally be considered skills problems. And in both cases the students resist the characterisation of the problem areas as skills issues which might be resolved through individual action - in Sue's case - or collaborative action - in Sao's case. The perception - or representation - of the problems as vocabulary may be a desire to identify them as knowledge-based, and thus resolvable through transmissive pedagogic action on the part of the teacher. This student perspective is examined further in Chapter 7.

Anna uses two teaching strategies when she explains vocabulary from the texts. Sometimes, perhaps when there is time or when the vocabulary items are ones she has identified as meriting attention, she engages students in a kind of Socratic dialogue where students are guided to understanding by comments and questions from Anna, e.g.:

Mari: Line 33 brasshat

Anna: Yes, that is an interesting expression, can you guess?

Several try:

Mari: Feminist?

Sara: South American clothing?

Anna indicates they are not correct, asks them to read again:

Anna: What is it associated with?

Sue: Soldiers.

Anna: Yes, that's all. It is a reference to the hats the military wear, so it means military regimes.

Classroom notes CS2:10

On other occasions Anna provides an explanation. The classroom notes of Week 7 provide an insight into the structure of these explanations:

Sue asks about pundits lampoon.

Anna explains, an explanation which has different elements - synonyms - which seem to be what students like - they all write the synonym in their text; Origin - pundit from a Hindi word; and context of use through examples - here the film series - National Lampoon which is not familiar to the students.

Anna: Anything else?

Classroom notes CS2:7

Students appreciate both the type of information on specific vocabulary items provided by Anna, and also the clarity with which it is given - it is a comment in the evaluation discussion which the students can easily agree on and several comment on the range of fields Anna can refer to in the interviews conducted as part of this study. The following long excerpt from an interview with Laure illustrates the different ways in which this type of vocabulary activity is satisfying for her.

Laure: Yesterday we did the text, Policing for Profit (23/4/2), and it was really interesting, because we first we have to read it by our own, and then, what I really appreciate, everyone said a word of vocabulary that he didn't know, and we could improve our vocabulary, and notice that in fact we have the same difficulties, because all the words that I didn't know, Sue or Della didn't know them either, so it was great to notice that we have quite the same level,

Int: About the words, this was interesting what you mentioned there, because this was one of the points dealing with vocabulary in the evaluation,

Laure: Yes, and I think that in the final examination, we will have to find some words, we will have an exercise about vocabulary, so in this kind of text it is good, because for example, we have many words that we don't know, perhaps if we make the effort to find the

general meaning, as she asks us to do, perhaps she can help us to understand the sentence, and then to try to find the definition of the meaning of the word, and I think we will have this kind of exercise for the final exam, so it is nice to do that before.

Int: Yes, it is. When you were reading this in the class, did you mark some words to ask about?

Laure: Yes I underlined some words I didn't know, and then in another colour I underlined the general ideas, because I thought that like last week, we will have to do an oral presentation, to find the arguments better when they are underlined.

Int: And what did you do as the teacher and the others were explaining?

Laure: Oh, I note them, I write them on a piece of paper, I can show you after, because I have got them, I am going to give them to Ina, she is going to finish her class at three, yes, I used to write all the words in red, and put the definition in English, I think I did that in English, in blue.

Int: And you feel that you will learn those words?

Laure: Yes, yes. There are some words that I knew, that, I don't remember which words, it was about, it (23/4/2) is not my text⁴, so it more difficult, for example, Della asked for some word last class, and I already knew it, and some other words, I don't remember what they were, but there were some words that I knew, but it's strange because for example, when I read the text I didn't really notice that one word, I didn't really know this word, but when Sue said some word, I didn't really pay attention to that, so it was good for me in fact for me, in fact, to do that all together, because sometimes when you read the text, you can read it and understand the general meaning, so you don't really pay attention, and you don't really make the effort to underline this word, but in fact, after that, you realise that if this word is alone you won't understand it, so I find that's really interesting,

Interview with Laure CS2:10

First, Laure can improve her vocabulary in this way, although she acknowledges that Anna's emphasis is on 'making an effort to find the general meaning'.

Second, she can tell how she is doing on the programme by comparing the words she knows or doesn't know with those of other students. Third, she is developing a specific examination skill. Fourth, she is building a database of new vocabulary, useful to her and to others. Fifth, her attention is drawn to words she doesn't realise she doesn't understand when a colleague asks about them.

This positive view is not completely unanimous. Dave, referring to the same episode as Laure, comments on the lack of benefit of this kind of classroom activity:

- Int: After the reading (23/4/2) you had a discussion of vocabulary and phrases, did you find that useful?
- Dave: I don't think, very useful, more of a curiosity for some words, you don't need to know every single word of the text to understand the text, even in your own language, you go through the papers, and you don't understand every single word, you can understand what they're saying, so, I don't think we should go through all the details of particular words, if you want to get, if you are curious about them, just underline the word, and look it up afterwards in the dictionary.
- Int: That's what the teacher was saying,
- Dave: Yes, but you know, she went on,
- Int: Why do you think you went on discussing individual words?
- Dave: Because some students kept on asking questions so she gave answers.

Interview with Dave CS2:9

Dave's explanation may be the best there is: Anna, similarly, does not think that this approach is a good use of classroom time:

- Int: An interesting aspect here was the opportunity you gave them for bringing up items of vocabulary.
- Anna: Yes, [...], that was one of the points that came up in the evaluation, and what is interesting about some of these things, is that now having had a history of teaching on this module for a couple of years, probably about, maybe four or five times. In the past I used to do a lot of vocabulary, and I stopped doing it, as a result of evaluation, people said we don't like doing this, so I stopped doing it, this particular group say they want it, so I try to do it for them.
- Int: It could actually have gone on for a long time
- Anna: Yes, and I used to do that, I used to make that a really important thing. What happened was, some people get bored with that, they don't want to spend the class time doing that, and it's almost like, "I don't want to listen to other people's things", kind of, something selfish.
- Int: What was interesting for me on this, watching them read, was that very few of them use dictionaries in the classroom, and it somehow surprises me, that they don't,
- Anna: I'm wondering whether I have sub-consciously developed a teaching style which, not stops them, but makes it seem less, maybe

somehow beside the point, because I have got a memory of people in the past using dictionaries all the time, and me feeling that that was getting in the way of understanding what they were doing, they were looking, very insecure, looking up every word virtually, but a lot of, I think I have developed a way of teaching which is that there are broad swathes of meaning, there are approaches to writing, and if you understand the approach, then you can understand, things you think you don't understand literally. For instance, in this Universities - Towers of Babble [sic] text, you couldn't use a dictionary and find out what Decline of Donnish Dominions means, the title of Halsey's book, you couldn't find it there, so, I don't know the answer to this, but it could be that I have developed a teaching style that takes attention away from the importance of individual words, as the key to understanding.

Interview with Anna CS2:6

Anna illustrates here a kind of internal dialogue regarding vocabulary, a process of reflection into which the evaluation data feeds and which explains the dissonance of thinking and action which is represented in Dave's comment above. The process of reflection and its role in both programme and teacher development are examined in the next section.

b) Programme improvement and teacher development

Underlying her varying practices in the classroom, there is Anna's fundamental principle, centred on 'broad swathes of meaning' in order to develop skill and confidence in dealing with academic texts. Anna acknowledges that she is uncertain on how to implement this principle, and is willing to change her teaching strategy, according to evaluation feedback from students. This change is not immediate: it seems to emerge after an initial expression of resistance and a period of reflection. Figure 5.9 below illustrates this cyclical process of developing a classroom strategy that realises her pedagogic principles.

Feedback		Resistance		Reflection		Innovation
CYCLE 1						
Anna gets feedback from students suggesting less attention to explaining words in class	→	Anna interprets this as selfish thinking on the part of students	→	Anna rationalises that this might not be the best way of using classroom time	→	Anna develops a pedagogy which focusses on comprehension of ideas rather than individual words
CYCLE 2						
Anna gets feedback from students suggesting more attention to explaining words in class	→	Anna resists suggestions that her focus on 'broad swathes of meaning' should change	→	Anna: 'This group say they want it, so I try to do it for them', and includes a short activity in Week 6	→	Anna spends more time on vocabulary in Weeks 9 & 10, and provides vocabulary tasks to texts in Weeks 11 & 12

Figure 5.9 Anna's response to evaluation feedback on vocabulary encountered in texts

This process of developmental action deriving from the evaluation activity, is supported by the description of the classroom activities in the classroom notes. The evaluation discussion took place in Week 5. The interview where the issue was still a key element of Anna's current thinking took place in Week 7, (actually four weeks later because of the Easter break). In Weeks 7 and 8 the texts were shorter, and there was no specific attention to individual vocabulary items. In Weeks 9 and 10, the texts are long and challenging, and Anna invites students to ask about words. Sue is first to ask in Week 9 in a dialogue which lasts twenty minutes and involves fourteen items, asked by six different students. The activity in Week 10 lasts twenty minutes also, with again fourteen items asked by five students. A further modification is evident in the materials in Weeks 11 and 12: the long texts have vocabulary questions at the end. It is as if Anna has decided that this feature of the sessions is here to stay, and she structures it as part of her planned activities, rather than something left to the response of individual students. Anna responds in a similarly hesitant and thoughtful manner to a related point raised in the evaluation - the length of texts for reading in the programme

(see 5.3b above). If this is a mechanism in teacher development, then it may explain the careful management of the procedure in the classroom. Anna had to manage the procedure in order to receive the comments and suggestions of students graciously, but without endorsing them or expressing any commitment to implementing them. Listening and noting were portrayed as key aspects of her role in the procedure. She also assumed the role of filter, quickly setting aside issues that were not for her to act on, such as timetabling, and exploring points where negotiated outcomes may be possible, as in the exchange with Sue quoted in 5.5a (page 144-145) above.

This analysis connects empirically two different aspects of English language programme theory: professional development of teachers, and the role of evaluation in the curriculum. These issues are particularly important for the management of programmes: they represent dimensions of the curriculum which are subject to intervention from bodies beyond the classroom, whether managers within the institution or professional bodies, such as BASELT and BALEAP, which assume responsibility for recognition and accreditation of programmes.

The nature of the professional development here - a form of experiential learning by teachers - conforms to models of teacher development described in 2.8 above. The process is especially resonant of the Schön (1983) form of reflective practice, i.e. where the process is embedded in professional practice, and is evident only in documentation of professional practice such as this thesis. The representation of reflective practice in language education is one akin to action research where a problem is explicitly identified and investigated, and a report subsequently published (Pollard and Tann 1987; Peck and Westgate (Eds.) 1994; Richards and Lockhart 1994; Wallace 1998). In these cases, reflective practice occurs in contexts of formal training - INSET and initial teacher training (see 2.8b above). In this study it is the research data that illustrates the process of professional development: it remains invisible as a dimension of the evaluation.

Woods (1996) describes a similar form of teacher change over time, where teachers' beliefs, assumptions and knowledge change with direct experience of

classrooms. The governing principle for this process is 'coherence' - an individual teacher's drive to bring thinking and practice into line in the enterprise of meeting the needs of a given classroom. Woods characterises coherence as involving:

the classroom actions and events, the teacher's interpretation of these actions and events, and the feedback of these interpretations into the ongoing planning and decision-making process that leads to further structuring in the carrying out of the coherence.

Woods (1996:249)

The study of the consequences of the discussion of vocabulary in the evaluation for both Anna and the programme, illustrates the explicit role of evaluation in this process. One might posit that this form of feedback is more likely to contribute to the teacher development process than feedback garnered implicitly on the classroom process, though these data do not allow such a comparative conclusion. The concept of coherence may explain the hesitant response of Anna set out above (and Millie - see 3.5c page 69 above): teaching action has to be based on a set of principles, and suggestions for change have to be reconciled with existing principles before they become available for action.

This professional development outcome of teachers' use of evaluation conforms to desired impact in the literature (Hopkins 1989; Rea-Dickins and Germaine 1992; Weir and Roberts 1994), though one which tends to be undocumented in published reports of evaluation. Evaluations of English language programmes which do not involve teachers, tend to recommend professional development activities rather than report such benefits as arising from the practice of evaluation (Mitchell *et al* 1981; Parkinson *et al* 1982; Lawrence 1993; Karavas-Doukas 1998). Evaluation initiatives which seek to get teachers to themselves document professional development benefits of evaluation practice may be unsuccessful (Brittan and O'Dwyer 1995), suggesting that institutions promoting evaluation for this purpose need to i) acknowledge the natural invisibility of such professional learning, and ii) select appropriate methods for accruing evidence of professional development where this is deemed necessary.

Professional development for teachers is a current priority across the higher education sector. It is considered a key mechanism underpinning quality *enhancement* of programmes. Chapter 6 explores how this desirable consequence of evaluation may be compromised by the quality *assurance* requirement, i.e. the reporting to managers and committees of evaluation findings. Chapter 8 revisits this problem, exploring how the latent potential of evaluation practice for teacher development might be realised and developed in an institutional context such as this.

5.6 Summary

This chapter has examined the validity of the evaluation conducted in CS2, using Messick's (1989; 1996) theoretical model as a broad framework, and the programme data as a platform for an *emic* description and explanation. The links between the assessment process itself and social consequences highlighted by Messick correspond to the central themes of this study: the social nature of the programme, the different stakeholder perspectives, the potentially conflicting purposes of the evaluation, and the notion of evaluation as a bridge between management of the programme within the classroom, and management at institutional level. These themes, examined in the classroom and interview data, inform on the different ways in which the evaluation of the programme is valid, and in which validity is threatened.

The following aspects of the programme evaluation attest to its validity:

- i) The points raised by students correspond to the classroom experience of the programme (5.2c and 5.3a above);
- ii) The NGT procedure facilitated negotiation of some aspects of the pedagogy (5.3a above);
- iii) The process produced a list which could be made available to managers and committees if requested (5.3a above);
- iv) The management of the NGT procedure by Anna enabled her to maintain and develop the coherence of her programme construct (5.3d above);

- v) The process stimulated teacher reflection and professional development (5.5 above).

Each of these facets of validity is, however, problematic:

- i) The three different programme constructs identified (5.2 above) make it difficult to focus the evaluation. Two separate problems are evident in the data: first, the students' understanding of Anna's construct of an EAP programme is limited (5.4 above), and thus their capacity to act as evaluation informants is compromised. Second, the NGT procedure, which is grounded in the experience of the students, may not, on its own, be capable of representing the complex weave of plans and intentions which this programme represents (5.2d above).
- ii) The negotiation of the remainder of the programme did not enhance the experience of the most disadvantaged students, and did not contribute to success in their main programmes (5.3c above). Rather, there seemed to be a silent awareness of their disadvantage, maintained both by Anna and the stronger students.
- iii) The report was not passed on. There are two separate issues here: first, Anna's commitment to the institutional policy for quality management (5.3b above), which is explored in detail in Chapter 6 below. Second, the report (i.e. the points listed on the OHP in 5.3a above) made no reference to the experience of struggling students, or Anna's apparent view (5.3c above) that they should not have been recruited onto her programme.
- iv) Anna managed the programme to avoid open conflict, and to maintain her control over the pedagogic strategies in the programme (5.3d above). While this feature of the evaluation may compensate for the fact that students may be poor informants (5.4

above), and may give Anna time to reflect on suggestions (5.5 b) above), it compromises the democratic, *emic* nature of the NGT.

- v) The professional development is a private, invisible process, evidenced in research data (5.5b above). The validity issue here is the degree to which this conforms to what the department and institution see as professional development. This issue is discussed further in Chapter 6 below.

This account of the evaluation of CS2 – a narrative with exploration of selected themes – derives from the research methodology outlined in Chapter 4 above, and research questions 1.1 - 1.4 (see 4.10 above). Space here precludes exploration of all the methodological and conceptual issues which were the subject of decisions at data collection and analysis stages. Two merit mention here: first the element of interpretation involved in analysing small pieces of data, and relating these to other data, and to constructs from the literature. For example, in 5.4 above the student responses to one particular classroom activity correspond neatly to Young's (1992) framework based on Habermas. It may be the theoretical strength of the Young framework which delivers such neatness. Or it may be a somewhat *etic* categorisation of students' comments: Lincoln and Guba's (1985) notions of *human instrument* and *tacit knowledge* (see 3.4 above) in conflict with the *grounded theory* of ethnography. The analysis can best be seen as tentative, a basis for ongoing inquiry into the complex weave of EAP and higher education classrooms more generally.

Second, the notion of consequences requires discussion⁵. Essentially, what happens after an event is not necessarily what follows from it in a causal sense. The account provided in this chapter of what happens after the evaluation event is a descriptive one: the discussion of links is a process of interpretation, supported as far as is possible by relevant data and theoretical constructs. For me, as human instrument, there is a coherence to the data-theory links developed in this chapter. This coherence is particularised (Stake 1995), embedded in this case and context. The extension to other cases and contexts is for teachers and researchers to

determine, to identify in what Woods (1996) labels a ‘dynamic notion of resonance’, where:

individual readers of the study process the resulting discussion according to their own interpretive processes, and look for their own coherence.

Woods 1996:46

The study of the evaluation event in this chapter has focussed on the teacher: her construction of the programme, her management of it (including the evaluation), and how the whole takes its place in her ongoing development of professional principles and strategies. The next chapter examines more closely the role of the institution, in particular the construction of its role as owner of the programme and the evaluation event.

Chapter 6

The role of the institution - sleeping partner or remote control?

6.1 Introduction

The discussion of the evaluation of the two programme in Chapters 3 and 5 illustrates the way that evaluation is the bridge between the management of the programme by the institution and management of the programme by the teacher within the classroom. The institution has two formal roles in the shaping of the programme: first it specifies aims and content; and second it establishes the evaluation policy, in part the NGT procedure conducted in Week 5. These functions realise the institution's commitment to quality and quality assurance, a commitment which is an element both of mission and of legal responsibility under the orders of governance. Despite this, it is clear from Chapter 5 that the programme is very much the construct of the teacher, Anna: it is her interpretation which determines the materials and learning activities, and permeates the teaching-learning discourse in the classroom. On the one hand the tension observed between teacher and institution is managed to the benefit of programme: involvement of the students, negotiation of the programme, and opportunities for teacher development. On the other, the divide in ownership may be a factor in the teacher's commitment to the evaluation, the sidelining of the student marginalisation issue (and thus feedback on institutional policy on recruitment), and the reporting of evaluation practice as evidence of quality assurance.

This chapter examines the tensions between institution and teacher roles in the management of the programme. It describes how, in this particular context, evaluation with its pedagogic and quality management functions renders this tension visible and researchable. Three sections explore hypotheses raised by the discussion in the preceding chapters. The first posits that the role of quality assurer on behalf of the institution is a problematic one for teachers. This is examined through an analysis of Anna's positioning on an institutional problem encountered by students. The second

hypothesis is that a policy of collaborative teacher monitoring and development of the programme for formal quality assurance purposes, as framed by managers, is problematic in terms of the work of teachers. The role of the EAP programme coordinator, constructed as a human bridge between management and pedagogic functions, is examined in this section. The third hypothesis is that these problems relate to recent developments in educational programme management, particularly in the higher education sector where both the construct of quality and methodologies of quality assurance are contested.

6.2 The institution in the programme - issues of positioning and representation

Three students on the programme - Mari, Sona and Sue - are exchange students, spending one year in the British institution as part of their home country university programme. In order to gain credits necessary for progression in their home university, they have to take an examination there in June, also the examination period in the British institution. One solution to this problem is for the students to return to their home university at this time, and for the British university to arrange to have examination papers for programmes such as EAP sent to that institution and administered as an examination there. Then papers are returned to the British university for marking in the normal way. Figure 6.1 summarises this episode over the last three weeks of the programme.

Week 9	Mari and Sona notice that the EAP examination date is after the date they plan to return to Spain. Anna advises them to discuss the problem with their programme leader.
Week 10	Mari and Sona tell Anna that the programme leader cannot help them. Anna advises them to go to the Examinations Unit, to ask about off-campus examinations. Sue states that she would like this arrangement also.
Week 12	Mari and Sona tell me that they have after difficult negotiation with administrators in both universities, resolved the problem. Sue states that she, as a private (non-Erasmus) student would have to pay. Anna advises her to speak to a named administrator in the Examinations Unit. Sue goes there during the break, and then tells Anna that she has to authorise the funding of the arrangement. Anna gives her the name of a manager in Anna's department to contact.

Figure 6.1 Summary of the off-campus examination episode

Anna talks about her efforts to sort out the issue in an interview after the session in Week 12:

- Int: With that student, Sue, what was the resolution of that problem with her exam in the end?
- Anna: There wasn't a resolution, what she was coming back to me with was: she had done a lot of work, getting phone numbers, all kinds things in her home institution, and at the end of it was being presented with more problems that she hadn't solved, at one point it was: 'who was going to pay for the cost involved in this?' And there seemed to be some other question about, for me, to say something or other was all right, and I declined to get involved in that, because I have no power to manipulate exam systems and regulations between this university and outside bodies, and I think it's a cop out for an exam unit to be saying to a lowly module lecturer that she has got to say something or other about somebody or other who is on her programme, and where the university is creating difficulties, and having difficulties, so, all I did in the end is, I gave her the phone numbers of [names of department managers], and, she is leaving at the weekend, if she has time, to go up and see them and present them with the problem and ask them to try to find a solution to whatever this problem is, to ask them to negotiate a solution with this [name of exams administrator], beyond that I am willing to do her an attendance certificate, which states what mark she actually gets in the oral and written assignments, and indicates that she was unable to do the exam part of it.
- Int: What about her other studies, her other exams?
- Anna: I just don't know, I just don't understand how some of these peoples' programmes work, because there is something in Germany she has to go back for, and I think that's to do with her studies in Germany, she has to return now, before our exam period actually starts.
- Int: So she is not coming back here for any exams, then?
- Anna: As far as I know she isn't, but I can't, it sounds as if she is sitting exams in Germany, that she has to go back for, and if she doesn't go, it causes her enormous problems in Germany, for continuing with her studies, so I don't know, [...] I've done my best, but I get to a full stop after a while (laughs)
- Int: Well, it's odd, that the other students, the Spanish ones, who can't do the exam here apparently have an arrangement to do it at their home institution?
- Anna: Well, they're on the ERASMUS programme, I think there is some help for them from the ERASMUS people, that's the impression I get, because they have got [name of international office administrator] as a contact, but this other girl, she seems to be doing it on her own.

This episode is telling in a number of ways:

a) *Boundary management in the classroom*

It shows Anna managing a boundary between her programme and the processes of other units within the institution. She positions herself on both sides of this boundary, i.e. as a member of the classroom community, a participant in the classroom and defender of students' interests on the one hand, and as a representative of the institution on the other. An example of this positioning is in her initial response to Mari and Sona's problem: she is a symbolic member of the institution in knowing the appropriate procedure, and advising students how to access it. Then, as she comments on the difficulties likely to be encountered she identifies with the classroom: the unit is a 'bureaucracy', it doesn't manage to solve problems 'students and lecturers' have. She gives an example of the rooming problem - location of groups in the appropriate sized classrooms. There are other instances in the data of this dual positioning in interacting with students, for example when Della asks about a certificate from the programme:

[during the break] Della returns, and asks Anna about a certificate of attendance - with marks on it. Anna says that she will organise this to the delight of Della. Then Anna reminds her that it will not be soon - because she will have to wait until marks are approved by the boards, and because of delays due to changes in organization and administration in the university.

Classroom notes CS2:8

Here Anna is responding positively to an administrative requirement, but with the proviso that it cannot happen as efficiently as Della might like, due to factors beyond the teacher's control. Similarly, casual references in the classroom discourse to factors which lessen the efficacy of the programme, such as the overhead projector in the classroom and the timetable, locate the source of the problems with non-teaching parts of the institution.

Helping students with programme-related problems, a predictable, traditional reflex for teachers, involves, in the case of the off-campus examination, Anna identifying with the institution. She does this by providing information, again a conventional teacherly function, but making clear that she has no executive authority to take action. She is prepared, as in the case of Mari and Sona, to inform students of what is possible, and who they should contact, and where they can find the offices. When this is insufficient, as in the case of Sue, she offers to get information from a personal contact in the institution (who would not normally be involved in the administration of Sue's programme). She does this by phone during the final session. When the information suggests a greater degree of involvement and commitment from Anna, she withdraws from the problem, leaving Sue with a manager to contact. The interview cited above shows that Anna is exercised, perhaps angry, at the suggestion from the examinations administrator, communicated through the student concerned, that Anna should assume responsibility for solving the problem. This anger may be due to the difficult position the situation places Anna in, vis-a-vis the student: she could not simultaneously be the defender and helper of students, and refusing to take the action that a seemingly authoritative body said she should.

Evaluation is also a context of boundary management, where Anna has twin goals: to conform to the institution-wide discourse of quality management through carrying out the recommended procedure and documenting the issues which arise on the one hand, and to manage the internal dynamic of her programme on the other. The former seems to involve achieving a minimum threshold - a list of points raised by students - while the latter has a more pedagogic function - the procedure is integrated into her teaching. In the plenary discussion (see 5.3a above) Anna deals publicly with matters relating to the needs of students in their main subject areas (the length of texts), strategies they might adopt (vocabulary notebooks) and the institution (timetabling; private study time). In the interview data, however, it is other issues which merit comment: the reading materials devised by Anna and the tension these generate

between students' needs (long texts) and wants (short texts); and problems due to the students' limited efforts to improve their English, and contribute to the classroom process (see 5.5 b) above). The boundary seems to be located between the 'public' discourses of the programme - issues which are clear to all participants - and the 'private' concerns of Anna - her own particular contributions to the programme, and how these might be fine-tuned. Although there is a possibility that this divide might derive from factors other than the role of the institution - for example the opportunity for reflection that Anna had between the classroom event and the interview - the conclusion that the boundary exists and is actively managed is a strong one. The implications for this in terms of institutional policy on evaluation are explored further in Chapter 8 below.

b) Role and status of teachers

The off-campus examination episode and the delicate boundary management outlined above, reveal a problem with the role and status of the teacher, both in relation to students and to administrators. First, for students, Anna is the key to the solution of their programme-related problems: they come to her in the first instance. As discussed above, Anna's response has two elements: to be helpful to the student in question, and to limit her role in the solution to information dissemination. This worked in the case of Mari and Sona's examination, but only when the students put in substantial effort. The case of Sue highlights the collaborative roles of teachers and administrators. When money is involved, as in the Sue's case, the issue of authorisation arises, and the problem returns to Anna. She does not feel that she has the authority to deal with this (she is a 'lowly module lecturer'), or that it is her responsibility to make the case. Her response is to give Sue a contact number for a manager, acknowledging, perhaps, that the solution was not within the gift of the administrator involved. The problems of dealing with money (and the buck-passing involved) in a traditionally public sector institution relates to the wider context discussed in section 4 of this chapter.

The role of the teacher, both in ELT and EAP, and in higher education in Britain has changed significantly in recent decades. The communicative approach and the democratic English language classroom have redefined the teacher as, *inter alia*, advisor, facilitator, organizer of learning resources, as well as instructor (Wright 1987; Legutke & Thomas 1991). In the context of higher education (where there is the additional battery of functions related to research and scholarly activity - see 6.4 below), there has been a similar expansion of role (Gibbs 1992; Race 1994). McNay (1995) quotes from an article on customer orientation in small businesses to characterise the entrepreneurial values growing across the higher education sector through the nineties:

Delighting customers means continually coming up with something unusual, which takes the customer by surprise, and which makes your company (college) and its people stand out from the crowd. It is about understanding and anticipating their needs, constantly seeking out problems, and quickly solving them for the customer.

Gilliland 1993 cited in McNay (1995)

This new set of values is characterised by McNay (1995) as emerging in a context of changing expectations of higher education by both students and staff (see 6.4 below). The implication is that those in the front line in 'service delivery' i.e. teachers, have roles quite different from the traditional don or tutor. Anna might consider herself working with the role outline in the second part of the Gilliland quotation, but would probably find the 'delighting' function too much.

The practice of evaluation as represented in this study gives the teacher the role of feedback organiser, negotiator, reporter and, within the programme context, action taker. Where quality is characterised in terms of the student learning experience, and there is an explicit link between programme evaluation which is managed by the teacher and quality assurance, there is a very real possibility that the teacher's role becomes one of general troubleshooter. This underpins Anna's response to the issues described above in relation to off-campus examinations, and may be a more

fundamental problem in the viability of such an integrated approach to evaluation and quality assurance. Anna feels that if the teacher discovers or acknowledges a problem of a management or administrative nature, then the units within the institution will not resolve these problems - it will stay within the teacher's domain, resulting in her taking action to resolve it, or her credibility as a helpful teacher being compromised. This is the cleft stick position Anna finds herself in with Sue, and generates the anger evident in the interview data. This expansion of the role of the teacher through the evaluation policy is examined below insofar as it relates to the student role, and in Chapter 8, as it relates to the development of the evaluation policy of the institution.

c) *The role of students*

The episode informs on the role of students within the programme and the institution. In line with institutional development generally through the 90's, the university here described itself as 'student-driven' and 'student-led': a form of clientism, where the needs of students are considered central in strategic and operational management, administrative processes and academic activity (Green 1994). Thus, the examinations unit is not only required to administer examinations in a secure and fair way: it also has to meet the diverse demands of students such as those described above. An important dimension of quality is the actual response to student learning requirements (see 6.3 for a discussion of the evaluation policy as quality assurance). There is general recognition in the literature that the changes in higher education in the 1990s have resulted in a greater range of student profiles, with diverse needs, and that the task of institution is to facilitate not only learning, but also their personal and professional development and general well-being (Haselgrove 1994, Opacic 1994; Henry 1994). To achieve this, there is a trend towards student empowerment, framed in the discourses of the service industries, and the charter movement (the *Study Guide* for the EAP programmes, CS1 and CS2, has a students' charter section - see 3.5a above) which outlines student rights and responsibilities. There is a paradox at the heart of the discourse of student empowerment: it creates, in a context of growing and diverse recruitment, an awareness of problems and a demand for solutions without

providing the resources for responding to these. This is the case, for example, with issues such as overcrowding in laboratories and libraries (Green 1994; Haselgrove 1994). It is also the case with the off-campus examination episode described here. Meeting students' needs has become important for the examinations unit; even a limited number of students availing of options such as off-campus examinations, places overwhelming demands on resources of the unit. The problem flows back to the classroom and the teacher - frontline troubleshooter - who resists having this kind of issue become part of her programme. Thus, empowerment of students can be at the teacher's expense.

This affects the purpose and practice of evaluation: where the evaluation is constructed as teacher-owned and managed, but with the purpose of engaging with the issues which affect the students' learning, there is a danger of a trade-off between student and teacher empowerment. Thus, the role implicit for students in the evaluation policy may explain Anna's management of the event (see 5.3 above), and thus underpin her need for careful boundary management.

In conclusion, this episode illustrates the reality of operating such a stakeholder approach to evaluation and quality management. The three main stakeholders are engaged in an elaborate power play, where each seeks to coalesce pragmatically with another against the third.

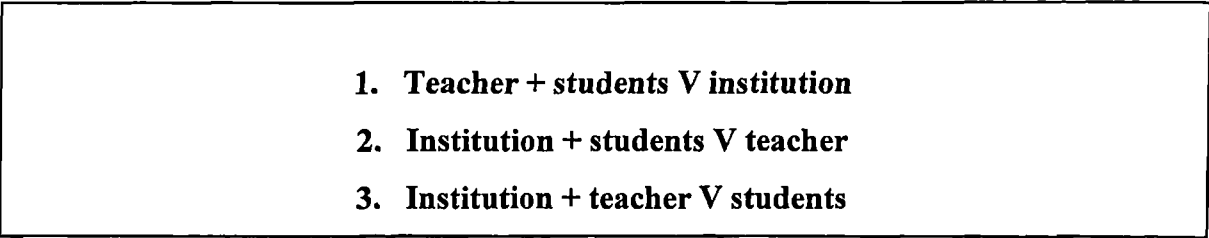
- 
- 1. Teacher + students V institution**
 - 2. Institution + students V teacher**
 - 3. Institution + teacher V students**

Figure 6.2 Stakeholder alliances

The analysis in this section supports hypotheses 1 and 2. The third might be considered a mechanism of an even deeper structure, with its roots in the traditionally elitist view of university education (which Anna's emphasis on the socialisation aspect of her programme might be considered an attempt to counteract). This hypothesis is explored further in Chapter 7, in the context of learning experience of an individual student.

The issues illustrated by the off-campus examination episode inform on the institutional backdrop of programme management and evaluation. The latter, the NGT procedure which Anna is implementing at the behest of the institution, contrasts with other elements of the programme which are her distinctly personal contribution. The tensions evident in Anna's dealing with the off-campus examination issues help understand Anna's tentative commitment to the evaluation described in Chapter 5 above. The two programme evaluations studied here - CS1 (Chapter 3) and CS2 (Chapter 5) illustrate the teachers' strategy for the task of boundary management: undertake the evaluation according to the policy set out, use it for the pedagogic development of the programme, and less consciously for personal professional development, keep a record of points arising in case they are asked for by managers or committees, but otherwise take no proactive role in reporting beyond the classroom. This strategy contributes to quality enhancement, but raises problems for the processes of quality assurance. These problems are examined in the following sections, as the role of the immediate organisational context - the EAP programmes unit - and the more remote institutional context are explored.

6.3 The programme in the institution - the inevitably gapped loop?

a) The customising of NGT

The evaluation policy background of the department within which the EAP programmes are located has been described in detail in 3.2 above. The discussion there illustrates the positive response of different external auditors and inspectors in the period immediately preceding the programme in question, and describes the NGT

procedure favoured for mid-programme evaluations (see also Appendix 6). This policy at departmental level is both an interpretation and implementation of the institution-wide policy (see 6.4 below), and a compromise between the agreed need for demonstrable procedures for quality management, and the need for teacher ownership and management of programmes - a form of collective boundary management.

The evaluation of the programme discussed in detail in 5.3a, above differs from the procedure set out in the departmental policy in a number of ways:

Procedure as set out in policy	Procedure represented in data
1. Two questions set	Three questions set
2. Period of silent individual work	Period of small group discussion
3. Master list of points from individuals	Master list from groups
4. Agreement on a final list	Agreement on final list
5. Indexing by students (Tally)	No indexing
6. Evaluative comment from teacher	No evaluative comment initially, but there is in response to points related to Question 2

Figure 6.3 Comparison of recommended procedure and actual evaluation

These differences are minor in terms of the conduct of the evaluation, but revealing in terms of Anna's management of the event, particularly in the context of the boundaries to be managed. Her third question, *Suggestions for the rest of the programme*, requires the students to develop solutions for problems they raise - a responsibility Anna holds them to in the classroom discussion - see 5.3 above. The small group discussion, especially with Anna's injunction that they agree on points, similarly limits the agenda. Although in practice, students were able to add points, the clarity with which Anna, as the teacher, had stated that agreement (and referred to it when receiving points in the plenary discussion: see for example her response to Sao, 5.3d, page 130 above) seems likely to have a constraining effect. The absence of a tally - a numerical value for each point - allowed Anna to manage the conclusions to the discussion, and her lack of evaluative comment in the early part of the procedure

could be read as a distancing strategy, a hesitation about engaging in the dialogue which might disrupt her boundary management.

b) Evaluation and programme development

The approach to evaluation sees the department as a forum for critical debate and consequential programme development (see Appendix 6). However, there is little evidence that this function is discharged in any formal way. For Anna meetings to discuss evaluation findings are a norm, something which should happen, but in the case of CS2: 'I don't think we have had a formal session, I think we have had informal conversations' (see 5.3b above, page 123-4 for a fuller quotation on this issue).

May, the EAP programmes coordinator, makes the same point about the formal operation of the policy on evaluation and quality management, and the informal discussions which take place:

- Int: Now, about the evaluations, do you get the evaluation findings from the teachers of all the groups?
- May: Yes. Not all of them. I ask each teacher on the module to make a report on the evaluation, and I collect those, officially those then go forward to the ELPG [English Language Programmes Group], I think only to the ELPG, but in actual fact, they probably stay on my file.
- Int: And what do you do with them? Do you talk about them with teachers?
- May: Well they get talked about with the students at the time, but beyond that, not until perhaps the end of semester, when we do a summative evaluation, and all being well, and time permitting, etc., etc., the summative evaluation tries to link in to the formative evaluation and see whether the semester has reflected everyone's concerted views in the end, and then we can discuss together, if we can manage to have a meeting of all the teachers, discuss any changes that might take place, and in fact, the present structure that we use in the second module [i.e. CS2] is based on agreements that we reached after a certain amount of experience and on the basis of those evaluations, but we haven't done it since, and that was probably a couple of years ago.

Interview with May CS2:9

The actual procedures of evaluation after the classroom procedure seem to be a blend of formal and informal encounters among the team. Formal processes seem to have a low priority for teachers: a report based on the NGT procedure can be constructed from teachers' notes if required. The conditions which May sees as needing to be satisfied before a meeting can be held illustrate the improbability of one actually taking place. May's coordinating role therefore has an administrative function rather than one which develops the curriculum:

- Int: OK, do all the tutors accept the need and see the value of a coordinating role for these modules?
- May: I don't know whether they all see the need of it, they certainly make use of the coordinator, I think, generally speaking, the coordinator can take away some of the housekeeping duties that a person otherwise has to think of for themselves, and it means that someone can ring you up, or send you an email and say what's to be done about this, and it's your job either to know the answer, or to find out the answer, so for the others it takes the little administrative burdens off the individual teachers, I think that's appreciated.
- Int: Would it be fair to conclude that the coordination has a lot of very useful administrative functions,
- May: Yes, I think that's true,
- Int: But isn't very strong on the actual curriculum functions?
- May: The curriculum functions, I think, are far more difficult, I think in the case of EAP, there was an attempt, I certainly attempted in the early days particularly, to look more rigidly at the consistency angle, but I realised very quickly what I should have known, probably, a priori, that teachers are fairly independent bodies normally, and they prefer to, each in their own way, though to keep, as I say a macro eye on what content is covered and so on.

Interview with May CS2:9

Two hypotheses emerge from these data. First, the reality of use of evaluation findings in this context is somewhat tokenistic: there is an awareness of what is required by the system, and activity is undertaken to provide this, without it intruding too much in the pedagogic enterprise in the classroom. This hypothesis sees the divide between management of the programme on the ground, and management of quality at the institutional level, between the agenda of the teacher, and the agenda of the

institution as ultimately unbridged. This is a common analysis in discussion of such evaluations in higher education. Opacic (1994) comments:

In reality, student feedback tends to be collected at the departmental level in a one-off, isolated and snap-shot manner, rarely analysed effectively on a year-by-year basis, let alone published or acted upon.

Opacic (1994:161)

The second hypothesis is that the mismatch is one of cycle as much as one of policy and agenda: teachers are not prepared to undertake a major review of the programme each time it is run, in the way which the evaluation policy suggests.

c) Evaluation and innovation - long and short cycles

The institutional approach to quality management is one characterised by a ritualistic, repeated cycle of activity, where procedures and events centring on evaluation take place in much the same way for each programme, every year. The teachers' approach seems to operate on a longer cycle: there is, at a given point, a surge of developmental activity, where the essentials of the programme, such as objectives, assessment procedures, and classroom activities consistent with these, are resolved by the team. The resolution is a framework with which each teacher can work. In CS2, the framework is 'the present structure' agreed 'a couple of years ago' (see May's comment in 6.3b above).

There is further evidence to support the cycle mismatch hypothesis. In commenting on the assessment procedures and criteria, Anna comments:

Anna: ... I am always concerned about the categories that I am going to assess, I don't think, I don't have full confidence in them, but as a group, committee, I have gone along with them. The one level of anxiety, I suppose, is one that might never occur, that someone could be put down in marks unjustifiably,

Int: That's these categories here, (looking at assessment pro forma), [...] There aren't any clear traditional linguistic categories here, like accent, intelligibility, pronunciation, and perhaps grammar?

Anna: That's right, I think that might be my influence, you would have to check it, but I think there was once stuff like that, but I argued strongly against it, on the grounds that I have a fundamental problem with the whole area of academic pronunciation, I'm not convinced that it is a very serious category, or easy to disentangle at all, and yet, and also part of the ELT tradition.

Interview with Anna Week 12

Anna's 'influence' refers to a key phase of development activity for the programme in the recent past which establishes a framework which i) she sees as still problematic, and ii) influenced by her thinking. Overall, for Anna, this seems to be a viable framework, within a collegial culture, where there is shared ownership and responsibility among the teaching team. There are tensions - above, in relation to assessment criteria - but the perceived problem is not unfair to individual students, and so is tolerable.

This long cycle of development has an established place in programme management in higher education: the concept of quinquennial review of programmes, a basic mechanism, for example, in the CNAA (Council for National Academic Awards) procedures which supported a range of institutions (including that where this research is located) in the development of programmes through the 1980s, and are still evident in programme management processes. However, attention to action based on student feedback, together with what Fullan labels *intensification* - increased definition of curriculum, mandated textbooks, standardised tests aligned with curriculum, specification of teaching and administrative methods backed up by evaluation and monitoring (Fullan 1991:16)¹ - has meant that the five-year gap between processes of developmental activity is no longer sustainable.

The longer cycle, however, where a major collaborative effort establishes the 'direction', and subsequent changes are small scale, conforms to the innovation cycle proposed by Chambers (1995), based on Hollander's (1965) innovation theory in industrial contexts, and applied to educational programmes in Weick (1976) .

Chambers distinguishes between big and small innovations: a big innovation takes place when a critical mass of stakeholders share a dissatisfaction with the status quo and agree that a change in direction is required. Once this is in place, there will be a series of small innovations, with amending, refining and finally resuscitating functions to make the original big innovation work. When the general view is that the small scale innovations are insufficient (for any of a number of reasons), there is a context for a big innovation - a major overhaul, and establishment of a new direction.

Stoller's (1994) study of programme development through innovation identifies 'viability' as the primary element of the path to innovation in curriculum innovations, ahead of the 'dissatisfaction' factor, i.e. teachers tend to innovate where the innovation seems likely to work more than where there is dissatisfaction with the status quo. The manner in which Anna changed her pedagogic approach to vocabulary in the classroom, because she felt the new strategy would work better in 5.5b above, may show how 'viability' promotes innovation. Millie's outlook at the end of CS1 - 'going further in the same direction' - (see 3.5 c) above), can be related to her view of the programme as viable, despite student dissatisfaction. Since the hypothesis suggests that a change of direction is a rarer event than a measure to go further in the same direction, one would expect that a survey like Stoller's should find more evidence of teachers seeking to make their programmes viable than relating feelings of dissatisfaction to innovation.

This finding suggests a link between the teacher development process described in 5.5b above, and the cycle of innovation evident in this discussion: teachers are much of the time focussed on making the programme work - a viability focus. In this process, feelings of dissatisfaction are marginalised, set aside until they accumulate sufficiently to dominate the agenda, and lead to a major reshaping of the programme. This analysis also suggests that Anna may be distinguishing between desirable changes which she can implement - the strategy for dealing with vocabulary - and ones which she cannot - the assessment criteria. Dissatisfaction seems to have an

inherent benchmark: Anna can work with the assessment criteria she is unhappy with as long as they are not unfair to students.

This cycle, organic in terms of the work of teachers, and the energy and synergy resources available within a programme team, is, however at odds with the ideological thrust, discourses and processes of institutional quality management. The mismatch may explain the gapped loops in May's functions as EAP coordinator. She is both a teacher and an academic manager. The set of functions associated with the latter, shaped by her understanding of the imperatives of the former, becomes a focus on administrative tasks. Her usefulness is expressed in terms of relieving teachers of burdensome bureaucratic processes instead of providing pedagogic leadership, developing an understanding of shared problems, or coordinating a response to them. There is little sharing of curricular experiences, or critical debate; instead there is a view that teachers are working towards common ends in their own particular ways, and it is best to let them get on with it.

Anna and May have a direct involvement in the EAP programme. Over the duration of one programme, and the cycle of many programmes, the account of Anna's work in this study, attests at several levels to successful teaching, and to benefits in terms of student learning, especially as recognised by the students themselves. Taking 'quality' then as a characteristic of experience and perception - the expressions of satisfaction by students, and my own feelings of relevance and engagement as I observed the classroom process and spoke to the participants - there is a strong case for concluding that it is a high quality programme. One might equally state that the evaluation process contributed strongly to this quality, and processes of quality enhancement. There is, however, a central problem here for the institution - it needs evidence of quality and quality enhancement to discharge its responsibility for quality assurance. The task for the institution, therefore, is to achieve a balance between the measures and resources to enhance the quality of the student experience on the one hand, and to document processes and outcomes on the other. The next section explores the extent

to which this balance is achieved. It examines the institution's published policy on evaluation (a key element in its approach to quality assurance - see 3.2 above, and Appendix 6), and relates this to wider issues in higher education in the last decade.

6.4 The wider agenda - evaluation, quality assurance and the framing of the teacher-manager partnership

Sections 6.2 and 6.3 above discuss a process of policy interpretation and implementation. At the department level, the policy on evaluation, is itself an interpretation of a policy set out by the wider institution. This policy, in turn, derives from the institutional approach to quality and quality management, and the trends and discourses across the sector which influence this. This section examines the extent to which the problems with evaluation policy and practice at classroom and departmental levels derive from sections of this policy chain which are remote from the classroom. It looks first at the institutional policy on evaluation, and discusses its links to quality management. Then it considers the characterisations of quality which have developed in the higher education context and their impact on evaluation policy and practice within the institution where the programme studied here - CS1 and CS2 - are located. The aim here is not to initiate a study into organisational and management issues in higher education, not least because the empirical element of the study is firmly located in the classroom and its near administrative hinterland. Rather, it is to acknowledge the issues which the teachers studied feel influence their professional activity, particularly with regard to evaluation. The ethnographic data explored in this and preceding chapters suggest that the notions of culture and community which characterise the EAP team are not self-contained: they are parts of a wider network of discourses and practices, and need to be understood as such.

The hypotheses which have emerged from the analysis in the preceding sections, and which are related here to the wider context are:

- i) The approach to evaluation and quality management sets up the three-way divide - students, teachers and management - and exacerbates the conflict potential of their different interests and concerns.
- ii) The institutional involvement in evaluation makes the process more bureaucratic, augmenting the work of teachers without added value to teaching and learning.
- iii) The approach to evaluation which does not conform to the long cycle of programme development outlined above derives from policy in the wider sector.

a) The evaluation policy

The institution in question has a policy on evaluation set out separately from its quality management procedures, but designed to furnish base data for these (see Appendix 6). While the policy is well-informed in terms of current issues of evaluation theory and practice (the 'General Principles' section conforms to the issues discussed in Chapter 2 above), there are two problems which relate to the hypotheses outlined above. First, the policy seeks to marry the teachers' professional agenda and the task of managers responsible for quality across the institution. Second, it specifies tasks without due concern for resources. This section explores the extent to which these problems with the policy derive from current trends in quality management in the university sector.

The purposes of evaluation are stated as accountability and development, glossed in the second sentence in point 1.1:

Stakeholder evaluation is undertaken primarily for two purposes: accountability and development. The University is committed to use evaluation to inform decision making and to aid the development of an effective teaching/learning environment.

from Evaluation policy section of Quality Management Handbook

The emphasis on the development function, which has more immediate links to teaching and learning, and thus to the enterprise of teachers, resonates through Sections I - IV.D of the policy document. This emphasis is further supported by the essentially democratic thrust of the policy, evident in two ways: i) the scope of the evaluation covers all aspects of the university's activity, including 'Departmental management' and 'the Directorate'; and ii) the issues of method: focus and frequency of evaluation are devolved to those in direct contact with students and other stakeholders. Thus, the policy is attempting to deal with conventional power imbalances between teachers and managers in educational institutions, and suggests a collegial organisational structure and culture (Bottery 1992; Bush 1995). Section IV.E refers to a pro-forma questionnaire designed for use across the institution, and which 'may' be used. This represents a break with the flexible, negotiable thrust of the document so far: ownership of the evaluation process is contingent upon its being satisfactorily done, with the default option being an evaluation of a programme through questionnaire data from students, without reference to the teachers involved. Section V describes the information flows to follow evaluation processes - essentially a merging of reported issues and action points from individual programmes to groups of programmes, to department boards, and ultimately to the Academic Board of the institution which has a legal requirement to monitor quality. This flow of information represents the accountability function of the evaluation process. It is essentially a reporting practice, a form of quality control. The term used within the institution for this flow and formal receipt by relevant committees of information is, significantly, *Annual Programme Monitoring*.

This monitoring process impacts in three different ways on the practice of evaluation.

- i) It counters the teacher empowerment characteristics of the first part of the evaluation policy. This affects the ownership issue, with teachers seeing

the evaluation as a management intrusion into their programmes, which must, therefore, be carefully managed.

- ii) There is the bureaucratic task of preparing a report. One problem here is the workload one - the teacher has to invest time in a task which has no benefit to the programme, and which represents the teacher and her professional skills to her line manager and others. It is revealing that the evaluation policy states that evaluation practice is not part of the staff appraisal scheme, and that this point merits special mention at the outset (see Appendix 6 Section I.G). This uncertainty of use factor - the official statement that data is not to be used to appraise teachers, and teachers' view that it might be - is likely to increase the complexity of the report writing task as some care has to be taken to hedge and modify comments.
- iii) There is the increasing remoteness from the classroom in the committees which receive the reports - themselves progressively summarised and generalised - and feedback on the issues. This remoteness is a function both of domain and type of programme, and time - if the feedback loop were to be completed, the teacher or module team, would have moved on and have a new set of issues to deal with.

At departmental level, the NGT is the recommended procedure (see Appendix 6).

This approach to evaluation might be seen as a pragmatic bridge between the teacher agenda of quality enhancement, and the institutional requirement for quality control in three ways.

- i) The data is oral, and is gathered in a forum managed by the teacher, limiting the element of threat and interference in the teacher's programme.

- ii) The reporting can be straightforward, as in the case of CS1, where report is the list of points raised by students and written up on the OHP by Anna.
- iii) The idea of a local team, not a remote committee considering outcomes and deciding action reduces the importance of feedback loops from more remote committees, which will occur much later, if at all.

This compensation at departmental level is not fully successful, however. All four teachers in the EAP team who returned questionnaires on the views and experiences of evaluation (see Appendix 2 below and 3.3 above) distinguished between the pedagogical uses of evaluation, and the less desirable, management aspects, particularly reporting. Anna (filling in this questionnaire before I asked her to collaborate in the study of CS2) comments in response to the cue: *Views on the role of evaluation*:

It's beginning to become a bureaucratic one-way street in which management use negative student evaluations against lecturers, while ignoring positive evaluations or evaluations which lead to inconvenient conclusions - e.g. students demanding more tutorial time.

Anna in questionnaire on evaluation before CS1

Negative attitudes to teachers concerning evaluation are not new. Anna's view here, however, is a specific critique of how the practice of evaluation - the existence of evaluation data - facilitates a management strategy, and thus shapes teachers' views of institutional management. Evaluation may profile latent conflicts. As discussed in 6.2 above, the institution wants to improve teaching quality and is perceived as willing to use student evaluations towards this end. The teachers take a defensive position, positioning themselves with the students and representing the institution as problematic. It is unlikely that the divide here is a function of the evaluation policy alone: it is rather a case of a local instantiation of a wider problem. The next section considers the extent to which changes in higher education which relate to quality

management and organisational culture contribute to the architecture of interests and alliances which frame Anna's experience of the evaluation policy (see 5.3 above).

b) Quality management in higher education

CS1 and CS2 are located in a British University in the mid 1990s - a period of significant change across the higher education sector, and especially in the new universities. Green (1994) identifies four factors especially important in policy development in the preceding decade:

- i) Rapid expansion of student numbers against a backcloth of public expenditures worries;
- ii) The general quest for better public services;
- iii) Increasing competition within the educational 'market' for resources and students;
- iv) The tension between quality and efficiency.

Green 1994:5

These factors can be related to aspects of quality management and organisational culture which inform on the evaluation processes described in this study as a whole, and more specifically, on the hypotheses outlined at the beginning of this section.

i) Quality Management

Quality has become both a buzzword and key concept in educational management in the 1990's. In the context of higher education in Britain, this is evident in the bodies responsible for extra-institutional management of the sector - Higher Education Quality Council which conducts quality audits, the Teaching Quality Assessments (TQA) of the funding body, HEFCE, and later, the Quality Assurance Agency (QAA). Quality is a nebulous concept in the higher education sector, not least because it is the subject of two conflicting discourses. First, it represents what the central funding authority values, and can use for rewards and penalties within the operation of the funding formulae - what might be labelled an *etic*² approach. These are quantifiable performance indicators, such as programme completion rates; degree classifications; and employment status six months after graduation (Johnes and Taylor 1990:9). Second, there is an *emic* view of quality, characterised as 'fitness for purpose' (Green

1994:15, 113) - the degree to which different programmes in autonomous institutions with diverse missions achieve their stated aims and objectives. The 1994-2002 inspection framework for teaching set out by HEFCE - the funder of higher education in England - has six areas: curriculum; teaching, learning and assessment; student progression; student support and welfare; learning resources; and quality assurance. Performance in each area is assessed against the institutions aims and objectives.

A range of critiques focus on the tensions between the two discourses: Green (1994), for example, taking a managerial perspective, examines the difficulties in finding a synthesis between the two. Becher (1989), Goodlad (1995), Barnett (1994), Cuthbert (1996) and Chater (1998) all take issue with the dominance of the absolute performance indicator approach, seeing it as a quest for efficiency empowering central management within institutions at the expense of academic departments. The more relativist quest for quality is seen as empowering students as customers over teachers, constructed as service providers (Goodlad 1995; McNay 1995; Schuller 1995)

The analysis of the quality management function of programme evaluation in this chapter develops this critique of the relativist view. First, the approach assumes agreement on what quality means within the institution - that there is a programme with a set of aims and objectives and everyone subscribes to a shared interpretation of these. The context within the institution is not necessarily like this, it is not a unified team with a common enterprise. As the off-campus examination episode (see 6.2 above) shows, what the students and university management want is not what teachers and administrators can provide, or consider a priority. In CS1 (see 3.5b above), we see the teachers and the institution promoting an independent learning strategy which students do not seem to value at all. Second, the relativist approach is not experienced as such by teachers and small teams; it is compromised by the potentially intrusive role of centralised organisational structures. The following section traces this development within higher education institutions, and links it to issues with evaluation policy and practice.

ii) Organisational culture

Anna's reservations about reporting evaluation findings (see 5.3d above) and the use of programme evaluation findings (6.4a above) seem to derive from her view of the prevailing culture rather than the articulated management strategy. The wider literature represents this culture in two ways. First, the experiences of academics is represented, by Chater (1998) as a fundamental change in the intellectual enterprise:

Knowledge is redefined as competence, creativity as income generation;
human resources are costs.

Chater 1998:232

Oakley (1999a) explores the nature of these changes in a novel: she describes

... familiar themes [where] over everything hangs the shadow of monetarism. Not much in universities is valued these days apart from what can readily be translated into financial currency. The old currency of scholarship has virtually passed away. But rather than actively mourn its passing, [people in universities] are sucked into a distracting game of acronyms, memos, mission statements, strategy groups, committees and endless unproductive meetings.

Oakley 1999b:20

Second, conventional management analyses such as McNay (1995) and Schuller (1995) illustrates how power has shifted in universities in recent decades from academics to managers, from the culture of the collegium to that of the corporation. McNay uses Weick's (1976) concept of loose coupling to distinguish between the looseness and tightness - largely autonomy and control - of the links between different units within an educational institution. In the collegium, both policy definition and control of implementation are loose, whereas in the corporation both are tight. The distinctions are represented along two axes as follows:

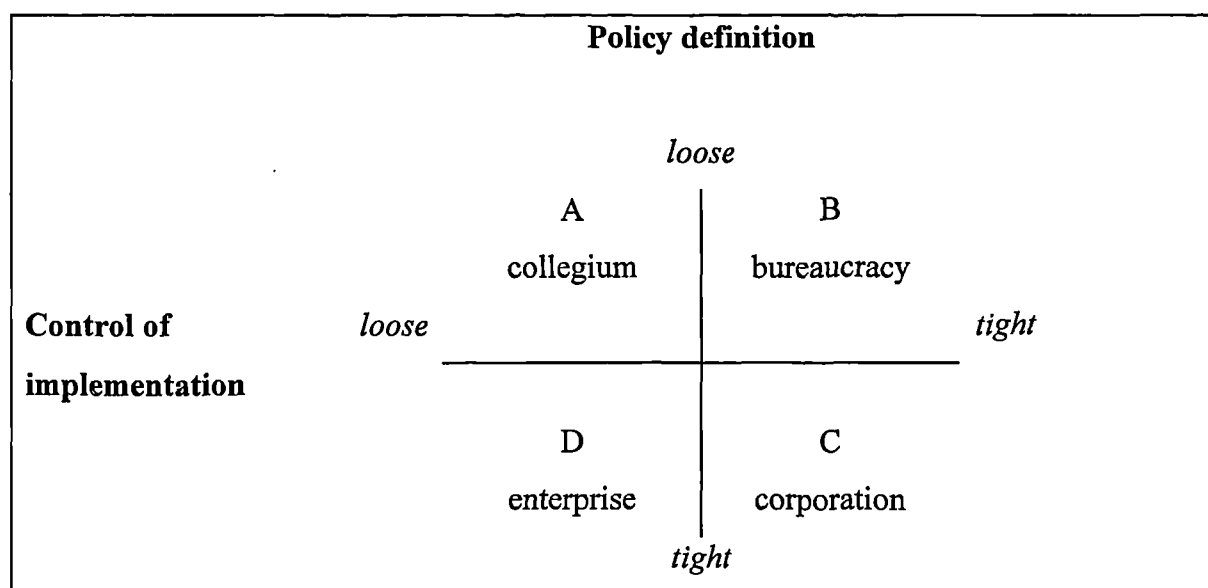


Figure 6.4 Models of universities as organisations (McNay 1995:106)

This overview provides a useful frame to map the links between the episodes and issues discussed in this and the preceding chapter. Anna works largely with the culture of the collegium, both in her construction of the programme (5.2b above), her management of the evaluation event (5.3d above), her response to students needs (6.2 above) and her own professional development (5.5b above). Her overall approach to the programme might be located in Quadrant A of Figure 6.3: a sense of working with the values of the collegium, characterised by teacher autonomy in respect of policy definition and control of implementation. This analysis is supported by May's perspective (see 6.3b above), who recognizes Anna's perspective as typical of a teacher. Her coordinating task is to bridge this and the bureaucratic features of academic management within the institution.

This analysis informs usefully on the three hypotheses outlined at the beginning of this section. The forces of organisational and cultural transformation outlined as part of the wider agenda account for the perceptions of different interests and conflicts, particularly evident in relation to teachers and managers. As Schuller (1995:6) states:

... a combination of broad economic trends and particular government ideologies has driven along a shift from administration as keeping things ticking over to management as active, proactive and even pre-empting professional practice.

Schuller (1995:6)

With management assuming the more pro-active role in determining how the institution operates, teachers take a reactive, even reactionary position, re-articulating the axioms of the collegium.

The second hypothesis - that institutional involvement in evaluation makes the process more bureaucratic, and thus augments the work of teachers without added value to teaching and learning - is both real and phantom. Producing a report for CS2 is likely to be arduous and time-consuming, where teachers like Anna feel such reports have hidden as well as stated uses. However, this does not have to be the case: the report produced after the NGT evaluation in CS1 (see 3.5b above) is largely a listing of the points generated by the discussion in the classroom. The workload issue may derive from the quest for efficiency (and associated teacher-manager conflicts) which characterises the wider context (see Green's four factors page 179): teachers on the one hand, are always likely to maximise their workload, and on the other, accord low priority to tasks associated with institutional bureaucracy.

The hypothesis concerning the different cycles of programme development is more complicated. There is a tension between the evaluation, student-led approach to quality, and the stability of an institutionally-approved programme. The process of approval of programmes within the institution in question, inherited from the CNAAP procedures, works on a five-year cycle - a programme is approved, a license is given for its delivery in that form for a period of five years, then there is a quinquennial review, which revises the programme as necessary. This approach to quality places a high value on the programme planning process. As discussed above, however, there has been a shift from quality as primarily a feature of planning, to quality as a feature of process. The latter sets up a situation where, if the students require it, elements of

the programme can, indeed should, change. While the quality management enterprise seeks to balance these two programme constructs, it may be ignoring the most important construct - the teacher's interpretation and management in the classroom (see 5.2 and 5.3 above). As the central mechanism for quality management - evaluation - thus misses the point, the thrust of the evaluation policy is to marginalise, rather than empower, the teachers.

6.5 Summary

This chapter, exploring RQ 1.6 (see 4.10 above) has traced the factors shaping a minor classroom episode to a range of discourses in the wider higher education sector. Along the way it has identified the perspectives on the institution of the teacher, Anna, and an academic coordinator, May. These perspectives share an understanding of how teachers work, and in relation to programme evaluation, show how the contribution of the institution needs to be managed. The departmental approach represents a synthesis of teaching and management cultures. The uncertainties and hedges evident at these levels of the institution are seen as having their sources in the wider context, in the transformed and transforming discourses of higher education in the 1990s, especially those of quality and evaluation. The data on one programme - field notes from the classroom, interviews and documents - thus frame a profound critique of the use of evaluation for quality management purposes. Chapter 8 revisits these issues, examining their implications for evaluation policy and practice on the one hand, and the training and development of teachers on the other. Chapter 7 explores the student perspective on the programme, illustrating how the discourses of the wider context (Chapter 6), and Anna's construct of the programme and evaluation (Chapter 5) impact on one particular learning experience.

Chapter 7

‘In fact, I can’t really lose’ - the case of Laure

7.1 Introduction

The focus in this chapter is the student experience. This complements the discussion in Chapter 5 which centred on the teacher as stakeholder in the programme and its evaluation, and Chapter 6, which explored the role of the wider institution. This chapter is a student case study, exploring the nature of the programme for a particular student, Laure, through two specific questions:

- i) What is the student experience of the programme (and how does this experience relate to different programme constructs), and the evaluation?
- ii) How do the evaluation functions - needs analysis, negotiation of programme emphases, etc. - benefit the student?

To explore these questions the chapter ‘tells the story’ (Stake 1995) of Laure’s experience of the programme, rather than the programme’s impact on Laure. Kushner characterises this approach to case study in evaluation:

Typically, evaluation reads private lives and personal ambitions through the spectacles of official definitions of programmes and goals. [...] we typically speak of managers, nurses, officers, teachers and so on treating people as creatures of programmes and institutions. We rarely identify the person behind the role - the lover, the parent, the hobbyist, the long distance runner. We rarely see programmes as vehicles or obstacles to personal realisation; more often we see individuals as vehicles or obstacles to the success of the programme.

Kushner 1996:194

Laure is a student majoring in English and British Area Studies (Spanish is her minor subject) on exchange from her French university for one year. She was one of nine

such students on this particular programme. As part of her languages study, however, she was taking courses in Law and Economics, as an introduction to the European context of language use. The focus of the English for Academic Purposes class resonated strongly with her sense of what she needed to achieve in her study of English, and how these needs might be realised:

Laure: I chose this course because [...] academic writing is the most interesting to really improve my English, and what I really like in that module is that we will learn how to make a good essay, to learn the link words and how to illustrate and everything, so that I why I chose this module, and also to improve my writing and my grammar and everything.

.....

Int: The activities in the class yesterday, did you think they were appropriate for you?

Laure: Yes, because as we are working in groups, it allows me to know the other students in the group, to exchange our ideas, to speak in English and to participate. Sometimes in a module, in Law for example, or in economics you only have to write, and in fact you stay alone during those classes, but in this kind of lecture you can meet other people. No, it was great.

Interview with Laure CS2:2

These glimpses of Laure's thinking regarding the programme provide a focus for the three aspects of her experience explored in this chapter:

- i) Laure's construct of the programme, in particular her emphasis on the social dimension of the programme and how this relates to the institutional and teacher constructs examined in 5.2 above;
- ii) The materials she likes and doesn't like: how these inform on her view of language learning, and on Anna's view of academic socialisation (see 5.4 above) and the awareness function of evaluation within the programme;
- iii) The negotiation of her essay with Anna is an episode which informs on the Institution + Teacher V Student stakeholder alliance hypothesis (see Figure 6.2, page 165 above) discussed in 6.2c above.

a) *Why Laure?*

The focus on Laure for this study has both conceptual and methodological rationales. First, she attended the programme regularly - 11 out of 12 sessions - and found the learning activities broadly relevant. Thus, from the perspective of participation, she might be considered a student who would benefit from the programme. Second, she provided me with a volume of research data: classroom observation data because she attended so regularly, recorded evaluation discussion data (discussed in 5.3a above); approximately six hours of interview data, including one hour of joint interview with Ina, her friend on the programme; and a written self evaluation / evaluation of the programme, which I requested after the formal evaluation event in Week 5 (attached as Appendix 7). Laure was a keen and enthusiastic interviewee, who felt she benefited from the experience. At the end of interviews I frequently asked her if she found it useful to talk to me about her learning on this programme, and her responses were generally positive. Her elaborations on initial affirmative responses varied, but were revealing of Laure's view of her participation in the programme and language learning. She seems to see the interviews as providing opportunities for learning in much the same way as programme input does. Second, her focus on the programme is on the process in the classroom, rather than any outcome which might be related to learning. These perceptions are especially evident in the first two interviews:

- Int: Do you like talking about your learning? Is it useful for you to talk about the problems you have in your study?
 Laure: Yes. It helps me to further improve, not to do the same mistakes again.

Interview with Laure CS2:2

- Int: Have these interviews changed the way you see the course? This is the second time I am asking you about these activities. Has it changed the way you think about this module?
 Laure: No, I don't think so, because last week I told you that I like this module and this week it was the same, because I really love this text [26/2/3 - text on the ethnic conflicts in the Balkans]. I really like that we are speaking with some other students to speak with them about these ethnic problems.

Interview with Laure CS2:3

Later, Laure is even more specific on the benefits of the interviews to the specific learning focus of the classroom activities:

Int: One last question, Laure, these interviews that I have with you during the module, is it helping you in any way with your academic English course?

Laure: Yes, because of the effort to remember what we did in the class, so in fact, I can't really lose, I can't really forget it, so yes, it's OK.

Interview with Laure CS2:9

In the joint interview with Ina after the final session, Laure's response to the interview process illustrates the direct nature of the pedagogic benefit, contrasting with Ina's view of indirect benefits to learning through reflection:

Int: OK One last question, then, Talking to me during this module, Laure, you have talked to me a number of times, have these interviews had any effect of your own perceptions of learning?

Laure: Oh, I don't know, it is true that, for example when I was speaking with you, for example one topic, sometimes I didn't really pay any attention to these things, because I thought it was not important, and when you were underlining that to me, as you were speaking about that, then I paid attention to that, and perhaps I wanted to read the text a bit more, perhaps to know it a bit better, but I don't really know.

Ina: For me, I really like the interviews (all laugh), I hope you have given more, it's true, I hope it helps you because it helps me to think about it, to take into account aspects you maybe don't pay attention to, but we underline to you.

Interview with Laure and Ina CS2:12

Laure's view of learning, particularly the emphasis on remembering, is explored in the following sections of this chapter. The excerpts from interviews above illustrates, in addition to her views of the programme, ways in which Laure contributed to the ethnographic orientation of the interviews: as I tried to probe certain aspects of her experience, she would talk fluently around this topic, or after touching on it fleetingly, would go on into other areas. She thus provided, in the somewhat vague, tangential quality of her comments on specific points, a clear account of her perspective on the programme, and the values and cultural norms which inform this.

7.2 The social dimensions of learning in the programme

Laure joined the programme in Week 2 - she chose it as an 'option' module, after discussions with friends who had done the module with Anna the preceding semester. There are two dimensions to her rationale for this choice, which might be labelled product and process reasons, as evident in her comments at the start of the first interview, cited in 7.1a above. The product dimension of the rationale is evident in her interest in 'making a good essay', the 'link words' and 'grammar', while the process element is represented by 'working in groups', 'knowing other students in the group', 'exchanging ideas, speaking in English', and 'participating'. The former is what she envisages taking from the programme, knowledge which can be consciously transferred to other study areas, while the latter connects with a more personal learning agenda: enhanced skills in English for communication purposes. Knowledge is associated with effort and hard work, traditional elements of study, while the skills element is associated with social, more enjoyable activities. This binary perspective on what she needs and what is benefiting her is evident in her written evaluation (see Appendix 7) and in the recorded small group discussion (see 5.3 above). In the latter, Laure, for example, proposes more talking in groups, but less in front of the whole class - a point the more industrious Della disagrees with.

The social aspect of classroom activities overlaps with life beyond the classroom for Laure. Overseas students in universities are often socially isolated (Lewins 1990; Ackers 1997; Furnham 1997), and see the programme and classroom as a key opportunity to make friends. Laure was strikingly enterprising here: she had become friends with Matti, a Spanish student, in the previous semester (who attended the first session of the programme in question), and through her came to know the Mexican students on the module. She explains:

Laure: [...] For example, when you know people, sometimes you don't really know what to say but when you have them in your class, after that you can discuss about the lectures, you have many oral presentations and so on to do at home and you can do that two or three, and it gives us also

conversation, but not a basic conversation about the weather or something like that, something more interesting.

Int: So you know the Mexican students...

Laure: Yes, you know, before I know them because my friend Matti is friends with them, but I only speak with them 'How are you?' 'Are you fine?' 'That's OK'. But now that I am with them in the lectures we are speaking about the lectures and we are walking together and I think it helps us to have a better conversation.

Interview with Laure CS2:2

The 'better conversation' included helping her out with using email to stay in touch with friends in other universities (though she stated later that the Mexicans did not know much more about email than she did). She got to know Della, and invited her to her birthday party in Week 9 of the programme. She had a conversation with Annie on a bus; advised Helen on cheap ways of getting to France for a weekend, and in turn received from Helen used telephone cards for her (Laure's) brother's collection. More than any other student interviewed, Laure tended to know classmates' names, and various details about them, such as their main study areas, and where they lived.

Bochner *et al.* (1977), in a seminal study of overseas students' friendship patterns at the University of Hawaii, identified three social networks in overseas students lives.

These are:

- i) A primary monocultural network, consisting of close friendships with other sojourning compatriots;
- ii) A secondary bi-cultural network consisting of bonds between the overseas student and significant host nationals such as academics, students, advisors and officials;
- iii) A third multi-cultural network of friends and acquaintances, of which the main function is to provide companionship for recreational, and non-task oriented activities.

Bochner *et al.* 1977:292, cited in Furnham 1997:19)

Laure eschews a reliance on i), on the grounds that she has not come to Britain to be with French people. However, as the programme develops, she becomes closer to both Helen and Arnie, the other French students. The bi-cultural network includes her landlady, with whom she became very friendly, and possibly me, but not Anna, of whom Laure always remained somewhat in awe - see 7.4 below. The multi-cultural

network is particularly important for Laure: she has a natural affinity with Spanish-speaking students because of her Spanish studies and her interest in maintaining her fluency in that language. She also associates easily with all European students in the class, compared to the Asian students, whose English she has difficulty in understanding. (No doubt, this unreflected view of Laure's contributes in part to the marginalisation which Asian students feel, as articulated by Niki - see 5.3 c) above.)

The social reality of the programme has an important status element for Laure: she rates herself against the students in the class whom she considers as 'strong', using classroom activities and episodes to inform this social dimension of the programme (see the interview excerpt discussed in 5.5a page 146-7 above). She categorises her vocabulary learning in terms of what she and strong students like Sue and Della did not know, and what she knew but Sue and Della did not. For Laure the learning activities and their relevance to her particular needs are realised socially in the classroom, as she works out who is doing OK and who is having problems, and how these phenomena locate her on the classroom ladder. Laure and Ina's discussion in the interview after the oral presentations illustrates not only their own interpretation of the assessment criteria being applied, but also their sense of how each individual student is ranked on this ladder.

This social element of the programme constitutes a key difference between Laure's construct of the programme and both Anna's perspective, and the institutional specification (see 5.2 above). Whereas the institution's formal account of the programme focusses on the functional element of the four skills, together with study skills and strategies for independent learning, and Anna's emphasises socialisation into the culture of British higher education, Laure's programme is peopled with potential friends, who constitute a classroom community, and are clues to how well she is doing. This programme construct contrasts with both her previous university studies, and other courses she is taking in the present institution, and represents the cutting edge of new experiences which are at the heart of the year abroad.

There has in recent years been increased attention to the social dimension of English Language programmes, both as a factor in enhancing the collaborative, communicative experiences of students (Breen 1985; Allwright and Bailey 1991; Legutke and Thomas 1991), and as a line of inquiry into how programmes work at a sociological level (Bailey 1996; Block 1996; Senior 1997a). The focus of such studies is usually the teacher, both as planner of programme activities, and leader in the classroom - Bailey (1996) and Senior (1997a) both derive lists of teacher strategies which determine the social and sociological structuring of the programme. This study, both the discussion of Laure's participation above, and the analysis of the social currents at play in the formal evaluation event (see 5.3d above), suggests that developing social networks among the students are significant in shaping the learning experience of individuals. The next chapter, echoing the concerns of Kushner (1996), examines the implications of this view of programmes for the practice of programme evaluation by teachers.

The twin elements of the programme for Laure - process and product - represent separate discourses. The social process is tangible, lived experience, blending classroom activities with life beyond, prioritising communication and people over linguistic and academic competencies. The product is an aspirational state of knowing, enthusiastically represented as about to be realised, but in reality likely to remain remote. The next section examines the links between these discourses, developing a hypothesis which sees Laure's particular understanding of how process relates to product as the source of the problem she has with the assessment (see 7.4 below). The questions this analysis raises for the evaluation purpose and process are discussed in the following sections, and taken up again in Chapter 8.

7.3 The materials - Laure's likes and dislikes

Laure is generally positive about the materials used in the classroom. This section examines her response to two kinds of academic English learning materials,

categorised as meta-EAP with a synthetic focus, i.e. attention to discrete linguistic elements of academic discourse (see 5.2c above and Appendix 5). The first kind is *link words* which appear again and again in the data provided by Laure as the focus of positive comment: she sees these lists of words and phrases as tangible products from classroom sessions which are the building blocks of the academic English skills she wants to develop. The second kind provides practice in academic referencing conventions, materials about which Laure is uncharacteristically negative.

Anna's rationale for the link words, and her account of the classroom procedure are discussed in 5.2d above. She sees these materials as working on different levels: the materials have a face validity insofar as they look like a formula for the problems of organising writing - an issue that resonates with Laure:

- Int: What about the other activities - you mentioned the link words [...] you can use to structure your writing, like "In addition, Actually"?
- Laure: Yes. I think it is really great because, in fact, when we are here we have to write a lot of essays, and the problem that sometimes you don't put the correct link word, and it is not really easy after that for the teacher to see the full meaning of the plan and everything. So I think that for that it is really great. I think it will help me to improve my writing, how to write more properly an essay ...

Interview with Laure CS2:2

- Laure: This one (pointing to 26/2/2) was really useful, we learnt the *transition*, all the link words, it is useful for the essay, I will use it for the next task. Because every week we have to write a task about some subject, and she [Anna] asks us to introduce these new kinds of link word, so it is useful to have them, ...

Interview with Laure CS2:3

Anna also shows a concern for the 'tangibility' factor - the notion of taking something from the lesson which builds into something complete over time, again a point that is reflected in Laure's experience:

- Laure: ... and [pointing to 26/2/2], whether I am going to read them again, because I read them before in the class, and I will read them again

when I will do my essay to try to introduce some words from this piece of paper to my work.

Int What will you do with this? Do you keep these in a file?

Laure: Yes I keep them in a, a file, [...] At the moment we don't have many pieces of paper, so I bring them with me to every lecture, but I think after some weeks, I will leave them at home.

Interview with Laure CS2:3

Anna, however, is somewhat hesitant on the transferability of the lists of link words to improved academic written discourse (see 5.2d above, page 116). Her comment here suggests that she would predict that the benefits to essay-writing skills which Laure expects 'within two weeks' are unlikely to happen. Anna would not be surprised at the real difficulties Laure experiences when she actually has to write an essay - problems described in 7.4 below.

Anna introduced two activities on the referencing and citation conventions in academic writing in Weeks 6 and 7. The activities are structured as a series of problem-solving tasks requiring students to integrate references and citations from published sources into samples of academic writing, and set out the bibliographical information appropriately. The tasks had a convergent structure (Duff 1986) i.e. students had to share existing knowledge about bibliographies and citation conventions to work out and agree on the correct solution. My own prediction on seeing the tasks was that students would respond positively:

It is a task which focusses on the technical aspects of academic writing, rather than on the textual or conceptual side, and feel that students should like it, should see it as really useful for them.

Classroom notes CS2:6

Several students - Ina, Arnie, Joe, Tigi, Sue - in interview commented positively on these activities, and there was a strong sense of engagement in the classroom as students first worked on the tasks in small groups, and then discussed the correct answers with Anna. Laure, however, proves a deviant case: she does not like the activities:

- Int: You did this one on the references (19/3/3)
- Laure: Yes, we did this one but I didn't really like this activity. You know, we don't really have those things, in France last year, we learnt how to put the references, I have in my memory how to do those things properly, so I already know how to do that, I didn't really know what to do with this one, because normally in academic purposes, we learn how to structure, we used to speak about the text, to give our point of view and understand, but it was really strange to do this, but, perhaps because we won't have the priority, others need this.
- Int: Did you have any problem answering these (19/3/3)?
- Laure: No, but sometimes we don't know how to say the words, like *hyphen* and *slash*.
- Int: What about this one (19/3/3 - second task)?
- Laure: Yes, we have to do that for next day, no, I didn't do it.
- Int: Will you do it?
- Laure: I think so, I will try, I will try, but there are so many things to think of, but I will try.

Interview with Laure CS2:6

The reasons Laure gives are varied: she has already covered this aspect of academic writing; it is not as important for her area of study as for others; it is not what she really likes and has come to expect in sessions; there were some difficult words in the tasks. When talking about the second activity on these conventions, however, Laure's thinking is more focussed - she does not see benefits to her writing, and she does not enjoy the process of doing these tasks which she finds difficult:

- Int: You did this one (9/4/2). Did you like this?
- Laure: It was to answer the question. Yes, we did this one in class, but I didn't really like it, I don't really like it when we are doing something like that, we did one about punctuation. [...] I prefer this text (9/4/3).
- Int: Did you do this (9/4/2) with somebody?
- Laure: With Helen, yes. I did it by myself before, and then with Helen. She [Anna] told us that we could compare with the other students, so with Helen, we tried to compare it. And this one (Task p.2) I tried to do it, but it was too difficult, because we have to put it in the correct way, but with that, this one (Task p.3), I didn't do it in fact, I was waiting for the correction, it was quite difficult, but this one, (Task p.1) where we did it with Helen, but some of our answers were wrong, but some were correct.
- Int: Was it new for you? The techniques involved here?
- Laure: I think yes, but I think before Easter we did something about punctuation, about the full-stop, and comma, all this kind of stuff, but

it's true, I think that it's a part of the module, but I prefer when for example, we have one text, and to speak in groups, or, even if I don't really like to speak in front of the class, I prefer it than doing something about punctuation, or at the beginning when she used to give us some link words, I find it really interesting, because it was really useful for our essay, and last time when she give us some information about referencing, first I didn't really understand why we had to do that, but after that I realised it will be helpful for our essay, but this one, I don't really understand why we had to do it, because even for our essay, I don't, I think I will write something about my own, so I can put the punctuation where I want.

Interview with Laure CS2:7

Laure's contrasting responses with these activities on academic writing conventions are instructive both about Laure as an individual student, and the evaluation of the programme more generally.

Flavell (1979) outlines three dimensions of metacognitive knowledge: person, task and strategic. Person knowledge is 'general knowledge learners have acquired about human factors that facilitate or inhibit learning' and 'specific knowledge about how these factors apply in their experience' (Wenden 1998:518). For Laure, a key element here is social interaction: learning is about discussion and exchange of ideas in English. Her belief system may also be informed by norms in foreign language study, where the referencing of elements of an argument is less important than the expression of such ideas. Task knowledge refers to 'what learners know about the purpose of the task, and how it will serve their language learning needs'; 'a classification process that determines the nature of a particular task'; and 'information about a task's demands' (Wenden 1998:518). The classroom notes and interview data show how Laure gets each of these wrong when doing activity 9/4/2. First, she and Helen, and later Ina, get into a discussion about terms for different punctuation marks (this group is next to me in the classroom, and my lurking and soaking is temporarily transformed to reluctantly informing!):

Students work silently, attentively.

11.30 Anna announces that students who want to discuss with a partner should do so - see if your answers are the same as theirs. Tells them to:

Anna: Do that if they think it is useful.

All students remain silent, then Laure and Helen discuss the task - Laure is trying to write an answer to each question, and asks Helen to clarify things - Helen at one stage asks me to help - it is about words for punctuation marks in English - hyphen; dash; and later exclamation mark: question mark.

[...]

Ina and Laure talk about the task - it seems that Ina has another view on what is a hyphen - Laure checks with Helen, who checks again with me - she says that Laure thought it was a dash. Then she asks me if a dash is /: I say that that is a slash. Helen then passes this on to Laure and Ina. Laure notes all these in her book.

Classroom observation CS2:6

This focus on punctuation fits with Laure's binary approach to learning activities: it is about essay-writing, like grammar or discourse markers. She sees the task's demands as declarative knowledge which she does not have, rather than reasoning from the data provided in the task. It is the punctuation terms which Laure refers to in the interview the following day, and remembers in the next session (three weeks later after the Easter break) when doing the second task on referencing conventions.

Strategic knowledge 'refers to general knowledge about what strategies are, why they are useful, and specific knowledge about when and how to use them' (Wenden 1998:519). The interview data suggest that Laure fails to take a strategic view when the task on first view does not seem relevant to her studies. She does not link the referencing task with the writing of her assessed essay (which she is doing at this time - see 7.4), or see the collaborative element as an opportunity for sharing ideas. Her perceptions of the task do not fit in with her view of her needs, preferences or expectations, and thus provide somewhat limited opportunities for learning. Thus, the interview excerpt (page 195-6 above) illustrates how her initial perception that the activity was solely about punctuation precludes her benefiting from it as an academic skills task¹.

Lea and Street (1998) (see also 5.4 above) outline three approaches to the development of academic writing skills in higher education:

- i) Study skills: a 'fix-it' approach to student deficit areas defined as atomised skills, and surface language features such as grammar and spelling;
- ii) Academic socialisation: acculturation of students into academic discourse, through a focus on orientation to learning and interpretation of learning task;
- iii) Academic literacies: student's negotiation of conflicting literacy practices, at the level of epistemologies and identities in institutions which are sites of and constituted in discourses and power.

Lea and Street 1998:172

These three approaches are not discrete or mutually exclusive; rather:

we would like to think that each model successfully encapsulates the other, so that the academic socialisation perspective takes account of study skills but includes them in the broader context of the acculturation processes [...] and likewise the academic literacies approach encapsulates the academic socialisation model, building on the insights developed there as well as the study skills view.

Lea and Street 1998:158

This model is helpful in understanding Laure's perspective: whereas the programme as set out by Anna is located between ii) and iii) (see 5.2 above), Laure sees learning as relating narrowly to i). This means that she does not appreciate the close link between her own valuing of social activities for learning within the programme, and the socialisation into the culture of an academic community which underpins the activities Anna introduces: she does not have a strategic perspective on the social activities she values, which might link these to the essay-writing skills which figure equally in her expectations. Wenden (1998: 520) suggests that Flavell's strategic knowledge might be considered a subset of task knowledge - it represents an ability to apply or realise what is understood about a learning activity in a given situation, to maximise the learning gains from a given learning opportunity. The data on Laure discussed here suggest that strategic knowledge may be considered an overarching category, embracing the sense of learner or student identity which underpins an academic literacy approach as outlined by Lea and Street (1998). The following section continues the study of Laure, examining this 'overarching strategy' hypothesis

in the context of the negotiation of her essay title with Anna: a process where Laure's view of the programme as set out above ultimately short circuits the achievement of her essay-writing goals.

7.4 Negotiating the essay title

The preceding section, at the risk of excessively representing Laure as an amalgam of deficits and shortcomings, might be seen as portraying a student who has commendable interpersonal skills, but lacks insight into the reasoning of others which is an element of negotiating skills. Dealing with the assessed essay supports this hypothesis, and illustrates the impact of the aspects of her experience of the programme explored in the preceding sections.

As part of the coursework assessment on the programme, students have to negotiate a title with Anna for an essay and a related oral presentation. This process can be presented in summary as follows:

Week 1	Assessment tasks pointed out in the introduction to the module, with students referred to Module Guide for details
Week 5	Anna reminds students that titles must be negotiated by Week 6
Week 6 (Last session before 2-week Easter break)	Anna briefly negotiates essay titles with all students except Laure: Mat, Joe, Tigi, Jamie, Sona, Sara, Sue, Mari and Rata at the start of the session while students are arriving; and Helen, Sao, Niki, Arnie, Della and Ina during the break
Week 7	During break negotiates title with Laure and Annie (absent in Week 6)
Week 8	Jamie gets extension; Essay due (Friday)
Weeks 11 & 12	Oral presentation based on essay

Figure 7.1 Schedule for coursework assessment

Laure was the only student in the session in Week 6 who did not agree an essay title with Anna. She talks about it in interview the following day:

Int: About the negotiation of the title, is this the first time you have negotiated a title like this?

Laure: Which title?

Int: The title for your assignment.

Laure: Yes, the problem is that my title is not really clear at the moment, in my head. I didn't talk about it

Int: Why not?

Laure: I wanted to speak about foreign languages, I needed to leave at 12 o'clock, and I didn't have time to speak before, I am going to France for two weeks, so I didn't talk about it. It is quite strange, because, in fact, we can write something about everything, so, I am quite afraid that my essay will be perhaps too general, but in fact, I said that I prefer when the teacher give us a subject of an assessment, and we can do it, and we are sure that we are doing it properly, because now, you know, with all the information and everything, so we can write something that is really general
[.....]

Int: You didn't want to talk to the teacher today to help you make the idea clear in your mind?

Laure: No, I didn't, no because I needed to leave at 12 o'clock, and the problem that during the break, I wanted to go to buy my ticket. And at the beginning of the class, the only people, but when it was my turn, I thought I will do it later, but then, I think I will talk about it after the holiday. Normally I will write something about foreign language, and the need for (unclear), because we can write about what we want, for example, Ina will write something about astrology, so I am not sure about what my topic will be.

Interview with Laure CS2:6

This account identifies lack of readiness and shortage of time as reasons for the negotiation taking place. Laure is aware that she has a topic, but not a focus or argument, and she knows this will be a problem. She identifies what the writing problems are likely to be, but does not see the negotiation with Anna as a means of overcoming these. In the following session, she speaks to Anna, and agrees what seems an even more general topic and title:

Int: OK. Can I ask you about your assignment. Did you negotiate with the teacher?

Laure: I am doing something about education. In the introduction I will speak about the history of the education, for example, the Education Act, I will speak about the Education Act of 1944, about the budget of the state, about the percentage of children who come to the private school, and afterwards I speak about the education, the necessity to be educated, to find a job, and also for our own, to know something, to

know foreign languages, to be literate, but my second part of it, there are some difficulties, so now I am blocked.

Int: But you have negotiated a title with the teacher?

Laure: Yes, yes, yes. First I wanted to speak about the education, the necessity to be educated, and about the education, and she told me that first it would be great perhaps to make differences between the education now and education in the past, to notice for example, she give me this example, in the past, in the farm, for example, the son used to have the farm, to exploit the farm as his father, but now, even if your father is a farmer, you have to go to a special school, to have a special diploma, to be able to do the accountancy, and to negotiate, and to know everything about the cultivation and everything, I tried also making arguments about that and for conserving the integrity of the education, but now, I'm blocked, of my second part, but I think that it will work, because, I wanted to, in my first part, to give all the arguments for education, and in the second part, against education, and I make also, a comparison with the education in France, in a short paragraph, but I have nearly a thousand words, and we have to write 1200 or 1500, so I think I will reach that for sure, but in the second part I don't have many arguments.

Interview with Laure CS2:7

Laure describes here the type of iterative process many students experience when working on an independent piece of work (Clark and Ivanic 1991). The concern about having enough to say to reach the word limit is especially resonant. What becomes clear at the end is that she has not quite agreed a title with Anna: she has spoken about education as a topic, Anna has tried to focus this interest with a practical example, but Laure is still struggling with how these disparate ideas - her own thoughts, Anna's comments, and issues from the literature - might fit together. She completes the essay on time, but is not happy with it:

Int: Tell me about your written assignment.

Laure: Yes, I wrote something about education, and it was quite difficult in fact, I regretted in fact, that I took this subject, because it's so wide, and I wanted to speak about everything, but the problem that perhaps I spoke about too many things, perhaps she will be quite confused, I don't know, but in my introduction I gave all the statistics about the budget, about the, how many children attend classes, and about the Education Act, and I gave the definition of education that I found in the Oxford dictionary, and I was speaking about the importance of

education now, because of the job, because of unemployment, of all this kind of thing, and then I spoke of the difference between education in the past and now, because I think that for example, when my parents used to go to school, they didn't learn like me, many things, like languages, or philosophy, this kind of stuff, but for example, they can write French better than me, or my brother, or the students from my generation, because they really learned how to write properly, and they really learned how to read also, properly, and I think perhaps they did less subjects, and they concentrated more on these subjects, so I spoke about that and I don't remember, and in the conclusion, I spoke about the new election, (UK general election May 1997) some changes will perhaps happen in the education.

Int: OK. Did you find in the activities and the guidelines you've had in the English for Academic Purposes module useful when you were writing this assignment?

Laure: The problem is that I wrote my essay without any notes, only helping me with the books and the dictionary, that I wanted to put the definition of education, but yes, when I typed it, I took the piece of paper about the link words to change, to really make, for example, as an introduction, I will say, or we, I think I used we, we will say, and as a conclusion, to really make a difference between introduction, conclusion, in the first part we will speak, firstly, secondly, or I said at first or at the beginning, and then I said moreover, and finally, and the same in the second part, I tried to organise my work, I tried, but I don't know if it is really, really good, because education is quite vast, and I wanted also to speak about, I spoke mainly about universities, but I also wanted to speak about nursery school and primary school, so I spoke quickly about that in the introduction, but then I did more my essay about education, but at the university.

Int: Good, so that was as you negotiated, as you discussed with the teacher?

Laure: Yes, but at the beginning I didn't know that we could choose about everything like that, I thought that we were quite obliged to write something about education, or languages, or about the field, our field of study, so I wanted to speak about education, but as at the moment, I'm writing my report (for another module) about teenage pregnancy, I would have preferred to work on teenage pregnancy, because I'm really more interested in that.

Interview with Laure Week 9

Laure got a mark of 42% for her essay. She did better in the oral presentation - 56% - and brought her aggregate mark for the programme up to 56% overall by getting 63 % in the examination.

This episode is telling in three ways. First, Laure had essay writing as a key learning goal for her, and felt through the programme that the activities were benefiting her. The analysis of her metacognitive awareness in 7.3 above suggests a lack of articulation - a strategic link - between underlying knowledge and operational skills with regard to academic writing. This is, at one level, an unresolved separation of programme process and programme product. At a deeper level, there seem to be conflicting dimensions to process for Laure: activities in the classroom are social, benefiting her by allowing her to interact with others, to exchange ideas and share perspectives. This process, however, is not transformative or developmental in terms of her academic voice, in shaping her academic identity, such that she feels at ease setting out a position and constructing arguments to clarify and defend this.

Second, when Laure was aware of a difficulty she seemed reluctant to engage with it openly. Perhaps as Abe in CS1 above (see 3.5b above), declined to discuss his plagiarism problems with Anna, Laure was not quite clear on the nature of her problem, and even more at sea on the shape of a solution, but she was not going to show these inadequacies. Her strategy for dealing with the problem was to avoid being labelled with it. Thus, she put off discussing a topic and title with Anna until she felt 'ready' or the time schedule obliged her to. Then, she clarified for Anna what she wanted to do, without really clarifying it for herself. It is as if the lists of discourse markers which she so values (see 7.3 above) are a metaphor: she has for her essay an array of words and phrases without a syntax or a discourse to make them meaningful. There is further evidence in the data of her awareness of how she is representing herself in the classroom: she was keen in interviews to explain what she considered poor performances in the classroom; and she also tended to liken herself in terms of needs and strengths to the 'strong' students in the class, for example, Della and Sue in relation to vocabulary knowledge discussed in 5.5a, page 146-7 above and 7.2 above.

Third, Laure is not overwhelmed by the difficulties she has with the essay. She remains positive about the programme and the range of learning opportunities it offers her (see her contribution in the NGT discussion in 5.3a above). A therapist, or tutor, might see this as denial, and conclude that unless she is brought to confront the demons of her unfocussed thinking, there can be no real progress. This, however, might leave her a depressed and alienated student who is not coping very well, as opposed to one bouncy in adversity. It is optimism for a purpose: the purpose of keeping her engaged, of sustaining her construction of a beneficial programme. She does not drop out, as Sao and Rata do, and being one of those who got through gives her an opportunity to continue her studies and feel part of the university community - representational evidence of socialisation into an academic community, even if the substance is still not in place.

The study of Laure set out in this chapter is a single case which illustrates how the programme becomes a learning construct, as opposed to the teaching construct of Anna, and the management construct of the institution. Describing and explicating learning constructs such as this is the task of evaluation, such that the student's awareness is raised, the teacher sees how her series of classroom activities are working, and the institution can assess the fitness for purpose of the programme. This study shows both the complexity of this task, and the degree to which teachers conducting the evaluation 'are in the dark' (Anna CS1:5 - see 3.5b, page 66 above). Two issues arising from this study: the potential of individual student case studies in programme evaluation; and the danger of using blunt tools for the evaluation task are explored further in the next chapter.

The purpose of individual case studies is not to derive some general principles or patterns related to the student experience of a programme - in many ways the study of Laure might be seen as constituting a negative case (Goetz and LeCompte 1981; Lincoln and Guba 1985), as in 7.4 where her particular experience appears contrary to trends within the programme. Rather, it is to explore the process of learning within

the programme, and the extent to which the planned programme of learning actually works. However, there are issues of research paradigm relating to this study as a whole, and to aspects of method in evaluation more widely, which the study of Laure problematises and clarifies.

7.5 A note on paradigm and method

This chapter started with a reference to the approach to evaluation of Kushner and Stake - leading theorists in educational evaluation. The reference is to the constructivism of their view of evaluation - the need for the evaluator to understand programmes from the perspectives of different participants, much as the ethnographer seeks to develop an *emic* account of a programme. Kushner and Stake, however, both work with what one might call a weak view of constructivism: a constructivist account is a partial one, not the full story of the programme, which is meaningful to a range of audiences. Kushner's view of programmes as 'sites of competitions between values and interests' (1996:193) reflects the need for evaluation to go beyond a constructivist account, and 'assert the possibility of a synthesising logic or analysis' (1996:198).

This account of Laure's experience of the programme builds on two constructions: hers, as the data reflect this, and mine. It is my own construction of Laure's experience that I wish to explore in this section. Reading through the classroom observation notes and listening to the interviews with Laure, I am a researcher with an ethnographic orientation. I am also aware that this is one voice in the Bakhtinian / Lemkean sense (Lemke 1995), overlapping in my notes and questions, with the voice of the teacher. And this teacher is one with an interest in evaluation, an activity which is about making value judgements, but only on the basis of full description and with a deep concern for perspicacity and fairness. Each voice derives from a social network, as Block (1998:152), citing Lemke to illustrate how his analysis of data from an individual student is framed, puts it:

We speak with the voices of our communities, and to the extent that we have individual voices, we fashion these out of the social voices already available to us, appropriating the words of others to speak a word of our own.

Lemke (1995:24-25)

This approach underlines the social dimension of utterances: the words used, the concepts alluded to, and the sequences of ideas, derives from practices within a social network. This assumption, significant in the analysis of Laure's view of the programme and her participation in it in the preceding sections of this chapter, is equally important in my perspective, as ethnographer, teacher, evaluator. The attention to the self is important in ethnography: the concept of reflexivity (4.8a above) requires the researcher to factor in an awareness of the discourses which inform the types of data collected, the content focus of these data, and the assumptions underpinning the constructs and hypotheses identified as emergent. (Goetz and LeCompte 1981; Lincoln and Guba 1985)

The reflexes of the teacher and the ethnographer are diametrically opposed in one respect: the teacher draws conclusions quickly, intuitively in real time, and while there may be subsequent reflection and accommodation of evidence which alters the initial view, that initial position often stands, and indeed becomes a filter through which later phenomena are perceived. My classroom notes record the first required individual performance by Laure in Week 3, when Anna is asking students to comment on a text about ethnic problems in the Former Soviet Union and Former Yugoslavia.

Next is Laure -

Anna: I caught your eye

Laure stands up, talks about the problem in Yugoslavia, between Serbs and Muslims, and how it is now an issue thanks to the international community. Gets increasingly nervous and flustered.

Anna: (helping out) What is the problem? What is the conflict about?
(Addressing whole class) Later in writing that is what you will need to address, what you will need to analyse.

Anna seems to shift from Laure who is not getting anywhere to a requirement for the whole group. Is this because she seems so flustered, not getting anywhere?

Classroom notes CS2:3

Anna chose this text for the specific purpose of getting students to take a fresh look at their own country:

Anna: ... one of the areas that was universal is the question of ethnic conflict, it occurred to me that everyone has some knowledge and experience of ethnic conflict in every country on earth so it is a suitable topic. Here was a quite thoughtful article related to real situations, ...

Interview with Anna CS2:4

Laure did not see it like this - she commented on Bosnia rather than France, because she had not picked up on this element of the task:

Int: When you were speaking afterward, you did not speak about France?

Laure: No. Because I thought we could speak about any problem in the world, so I spoke about the ethnic problem in Yugoslavia, because I found that quite strange, that two people in the same country, they can't reach to an agreement, to stop the war, and that everybody's killing themselves like that, no I didn't know it was about our problem, but if I knew that I will speak about the suburbs in Paris, between Parisian people and the immigrants, that there are still some problems between them. Because in Angers where I live we don't have many problems, there are not many immigrant people in Angers, so if there are some problems it is not really important in comparison with Paris. I will speak about, I would have spoken about Paris.

Interview with Laure CS2:3

Despite the activity not working out successfully for her in terms of the oral performance, or in relation to Anna's planned objectives, Laure comments positively both on the activity and Anna's dealing with it in class:

Laure: ... But here I found it quite interesting that people could express themselves and say their own point of view about that. I really like this module for that.

Int: Did you find the teacher's comments helpful?

Laure: Yes. Because she tried to in fact to tell us that, we have to develop more our ideas, how to structure, for example, I don't remember what she said to me, but for example, for Helen she said she should organise her ideas to say, for example, the beginning of the problem, to explain the problem, and at the end to try to say conclusion and to give her point of view. That's the same with the Spanish girl, Sona, that she should organise more, to say each idea, not to mix everything, in fact.

Interview with Laure CS2:3

My construction of this episode derives in five ways from my identity as a teacher. First, I perceive Laure's performance - nervous and flustered - in a teacherly judgemental way. Second, I account for Anna's response to what I see as the breakdown in Laure's performance in a teaching frame - using the instance as an opportunity to make a point generally relevant to the class - what Bailey labels 'sharing the wealth' (1996:24) - while at the same time, tactfully taking attention from Laure. Third, I share Anna's thinking on this kind of task - the pedagogic logic of getting students to read of one situation and relate to another situation with which they are familiar. Thus, I note Laure's *not* talking about such problems in France as significant, and ask her about it in interview. Fourth, I assume that the episode is salient for Anna as it is for me, and that it represents a defining moment in how Anna sees Laure: as a 'weak' student who has a lot of learning to develop the kind of academic voice which is the goal of the programme. Fifth, I hypothesise that Laure herself is aware that she has represented herself in this episode as a 'weak' student, but avoids a re-run of this part of the episode in the interview the following day, by focussing on Anna's feedback to other students. The data do not confirm these two last points: neither Anna nor Laure explicitly support these hypotheses in the interviews. The grounded theory can only be developed through a theoretical rather than data-based framework.

My response to the data here, however, is neither validating nor invalidating in itself: it is an aspect of the research process, much as the selection of the research problem, the particular programme and the phenomena recorded in the field notes, and reported in the interviews in response to my questions and cues. There are ethical issues for me

in interpreting the data, where someone, albeit anonymously, is represented in a negative way. Additionally, there is a theoretical issue. In Chapter 4 above, there is a discussion of the use of ethnography in educational research, and Green and Bloome's three approaches in particular: i) doing ethnography; ii) taking an ethnographic perspective; and iii) using ethnographic tools (Roberts 1998:56). At the outset of this research project, the role of ethnography related to ii) and iii). At this stage, after much analysis and explication of the data, the need to engage with the social and cultural theories, and related ontological and epistemological discourses highlights the need to regard ethnographic method as linking ultimately to i). The story of Laure, as indeed the other aspects of the programme which are characterised by their social and interactional features explored in this study, are only meaningful when related to theoretical frameworks in cultural anthropology. In this way, success can be related to survival and maintenance of a social identity and role, Laure's reflexes and actions to her interests.

This methodological and theoretical orientation is relevant to the practice of programme evaluation. The study of Laure's experience of the programme illustrates that a conventional means/ends account would represent the experience as a series of disappointed expectations and missed opportunities. The social perspective, developed from Laure's construct of the programme, illustrates how, despite problems of pedagogic engagement, the programme is a success for Laure. There may be a kind of poetic justice between the struggle which Laure has with the assessed essay, and the low mark she gets for it: it shows the coherence of Anna's programme construct, with its emphasis on academic literacies, rather than a set of discrete, mechanical moves. The main point of Laure's experience of the programme is that she stays with it, continuously engaged in the process, even though it is not working for her in an ideal way. Her own words "In fact I can't really lose" prove ultimately as sound a truth as any other. And it is an ethnographic account of Laure's experience, one which explains how the individual relates to the communal, which represents this.

7.6 Summary

The discussion in this chapter has sought to explore the ways the individual student - Laure - sees herself as a stakeholder in the programme (RQ 1.6 - see 4.10 above). Her stake is clearly framed in needs and expectations at the outset, although these constitute a different programme construct from that of Anna and that of the institutional specification - see 5.2 above. Laure's construct has social interaction at its core, and the programme supports the realisation of this. It does so, however in an individualised way: Laure interprets each activity according to her own discourses of EAP. An ethnographic perspective on these discourses illustrates their surface features and conceptual underpinnings. A pedagogic view outlines shortcomings and opportunities for development.

The second question in this chapter explores the extent to which the evaluation event provides a framework for Laure's participation in and benefit from the programme. Her participation as outlined in 5.3a above, conforms to Anna's requirements and to the participation of the other students. It is not transformative in any pedagogic sense: it does not raise awareness of the programme aims and objectives which enables Laure to develop a strategic perspective, or, more practically which facilitates her planning and writing a 'good' essay. Thus, the question of developmental benefit of the evaluation event to the programme, positive when explored from the perspective of Anna (see 5.2b above), remains somewhat undetermined in the case of Laure, the student who attended a higher proportion of the input sessions than any other.

The next chapter draws together some themes from this study of a programme and its evaluation. It links the focus on learning and the student role in the programme as explored in this chapter to the perspective of the institution and the teacher to elaborate some principles for the further study and development of programme evaluation and quality management.

Chapter 8

Conclusion: Rethinking, repositioning and rewiring evaluation

8.1 Introduction

The issues explored in this study - improving teaching and learning in higher education programmes; teacher development; developing language skills of overseas students; quality management within institutions and external quality assurance - are currently subjects of debate and policy development across the higher education sector in Britain. This research study, with its focus on evaluation and how these different issues intersect, informs on these debates in two ways. First, it shows how these issues are connected, for example, how one EAP teacher's response to the vocabulary issue (see 5.5 above) is shaped by pedagogic principle, student preferences, and institutional policy on quality management. Second, it shows how in practice, these connections are set aside, for example, how Anna is aware of but silent on the problems which arise from the recruitment of students like Sao and Rata onto her programme (see 5.3c above); and how May lets the evaluation reports sit on her file (see 6.3b above). This chapter explores five aspects of programme evaluation by teachers which relate to these broad themes of connectedness and dislocation:

- The different traditions of programme enquiry (8.2)
- Institutional policy and evaluation design (8.3)
- An ethical framework (8.4)
- Evaluation and teaching (8.5)
- Programme evaluation: towards an integrated management strategy (8.6)

These aspects of evaluation relate in a complex way to the research questions set out in 3.6c and 4.10 above, and to the different strands in the analysis and discussion chapters of this thesis. In this chapter I attempt to synthesise, without inappropriate reductionism or simplification, how the issues affecting programme evaluation by

teachers might be explored further. The goal is not to set out generalised recommendations: rather it is to show how the ethnographic case studies (CS1 and CS2) cohere, and invite readers to consider whether the discussion i) represents a way forward for this context, and ii) holds relevance for evaluation policy and practice in their context (see 4.8 and 5.6 above). Figure 8.1 summarises the findings of the case studies, and links them to the research questions, and to the different sections of this chapter.

Research questions	Key findings (Emergent hypotheses)	Discussion in this chapter
1.1 The nature of the programme evaluated	<ul style="list-style-type: none"> • Different stakeholders have different programme constructs (see 5.2 and 7.2 above) • The teacher's plan may be destabilised by evaluation and the teacher manages the evaluation with an awareness of this (see 5.3 above) • The programme for teachers represents a small part of a wider pedagogical construct (see 3.5c, 5.5b, and 6.3c above). 	8.2 8.5
1.2 The nature of the evaluation events	<ul style="list-style-type: none"> • The potential of the NGT procedure for engaging students in evaluation is real (see 5.3a above); • It can be compromised by i) the teacher's management strategy of NGT in the classroom (see 5.3d and 6.2a above); ii) problems with students as informants (see 5.4 above); • A key Quality Assurance function - reporting findings - seems incompatible with pedagogic functions (see 3.5b and 6.3 above). 	8.2 8.4 8.6
1.3 Characterisation of learning within the programme	<ul style="list-style-type: none"> • The evaluation focusses on teaching rather than on learning (see 5.3a above); • The notion of critical learning in the programme might be related to providing evaluation feedback (see 5.4 above); • Student lack of awareness of knowledge and skills issues in EAP results in vague feedback (see 3.5b, 5.4 and 7.3 above). 	8.3 8.5

1.4 The consequences of the evaluation	<ul style="list-style-type: none"> • There are benefits in terms of pedagogic focus and teacher development (see 5.5 above); • Strong students influence the focus of the programme post-NGT (5.3d, 5.5a) • Important issues are left unexplored (see 5.3c and 7.3 above) • Sectoral interests and stakeholders alliances are reinforced (see 6.2 and 7.4 above). 	8.3 8.3 8.6
1.5 The role of the institution	<ul style="list-style-type: none"> • The practice of evaluation generates quality enhancement but not necessarily quality assurance (see 5.5 and 6.4 above). • The democratic goals of the evaluation policy are not developed in the policy itself (see 6.4a above); • Teacher - institution collaboration is compromised by perceived sectoral interests (see 5.3b and 6.4b above); 	8.3 8.6
1.6 The experience of the student	<ul style="list-style-type: none"> • Students have their own programme construct, but do not negotiate a shared construct with the teacher (see 7.2 above); • Students participation in the programme and evaluation is guided by perceived strategic interests (see 3.5b, 5.3a, 5.3c, 6.3c and 7.3 above); • The NGT permits a co-constructed (by teacher and students) avoidance of key programme issues (see 5.3a); 5.3d above. 	8.4 8.5

Figure 8.1 Summary of research foci, key findings, and further exploration

8.2 The different traditions of programme enquiry

As outlined in 2.8 above such enquiry comes with different labels - programme evaluation, action research, reflective practice, classroom ethnography and teacher appraisal. While these have different epistemological and methodological origins, they all document programme processes, and share an account of these with others.

There are different purposes for this documenting and sharing: programme development, accountability, research, and quality assurance. In the context of CS1 and CS2 (Chapters 3 and 5 above) the institutional policy represents a broad commitment to these purposes (see 6.3b and Appendix 6). The departmental interpretation of this policy specifies a methodology of inquiry - NGT - but does not elaborate on the issue of purpose (see Appendix 6). In the classroom, Anna manages the NGT procedure, a received methodology, for programme management and development (see 5.3 and 5.8 above). This situation might benefit from a closer integration of purpose and method at each evaluation policy level. The departmental policy (see Appendix 6) might focus on the *why* question as well as the *how*. Teachers, the EAP team for example, could then engage with the NGT procedure with a stronger sense of ownership, and link it more explicitly to other forms of classroom inquiry.

In keeping with current concepts of the learning organisation, the institution needs to have inquiry informing all its processes, keying in to the existing culture of research in such institutions to document and share understandings of different facets of the programme processes. For example, there might be a meeting of the EAP teaching team to identify issues which merit further investigation, and develop hypotheses as precisely as possible from the existing findings. The purpose of such inquiry would be, broadly, programme improvement: for example, teacher issues such as materials evaluation; student issues such as language or learning awareness; institutional issues such as recruitment and placement. Figure 8.2 outlines how the issue of Asian student marginalisation – which had a silent role in CS2 (see 5.3c above) – might be taken forward through different inquiry strategies.

Issue	Purpose of study	Programme inquiry tradition	Conducted by...
e.g.	To determine the extent of the problem across institution	Annual programme monitoring (Evaluation / research)	Quality Enhancement Unit
Marginalisation of Asian students (see 5.3d above)	To document the experience of marginalisation	Action research (Case study on one student, e.g. Niki)	Teacher
	To document marginalisation phenomena in the classroom	Classroom ethnography	Researcher
	To examine student-student communication breakdowns	Small-scale Applied Linguistics research study	Teacher / Researcher
	To develop appropriate pedagogic strategies	Reflective practice Teacher appraisal	Teacher Human Resources Unit

Figure 8.2 Examples of different programme inquiry traditions in action

The initiation of such studies might be part of the remit of the evaluation development cycle discussed in 8.3 below. While space precludes a detailed discussion of all the issues involved in resourcing and managing such a series of studies, three points raised in preceding chapters merit mention. First, to alleviate the problems which derive from the academic-management divide (see 5.3d and 6.2a above), evaluation development requires collaboration between teachers and those more remote managers with a responsibility for quality. Second, the divide between evaluation and research (see 4.2 above) might be bridged by a more rigorous, academic approach to understanding the student perspective. Third, by relating such evaluation to Applied Linguistics issues, practice with English language programmes might incrementally

develop programme theory in the field. In an age of post-modern scepticism for grand theories and programme designs, improved practice is likely to develop through multi-faceted empirical scrutiny and fine-tuning. I turn now to how this might be achieved.

8.3 Institutional policy and evaluation design

The evaluation policy described in 3.2 and 6.4 above (see Appendix 6) has two levels: at a general level it outlines the merits and functions of inquiry into programme processes, while at department level, it specifies a design for the inquiry in each programme case. The outline policy derives from i) a legislated requirement for the institution to monitor the quality of its programmes, ii) a mission to be student-led, and thus develop its programmes through accommodation of students needs and preferences, and finally iii) a concern for a democratic, participative approach to management. The rationale for the department level policy - the adapted form of NGT described in 3.2, 5.3 and 6.3 above - has three principal elements: a grounding in the experience of students; a feasible workload for teachers, and convenience in communicating findings for the purposes of quality assurance. The analysis in this study shows i) that these elements of the rationale are not evident in practice (3.5 and 5.3 above), and ii) they are compromised by sectoral and cultural features of the higher education context (6.3 above). The following sections of this chapter examines the ethical, pedagogical and management dimensions of these issues in relation to the case studies set out in Chapter 3 and 5 above. This section focusses on the issue of evaluation design as part of the institution's quality management framework.

At policy level, the issue of responsibility for programme evaluation might be redefined such that use of findings and action-taking remain the responsibility of teachers and teaching teams, rather than Programmes Managers, Resources Managers and Departmental Boards (see Section V.A Appendix 6 below). Thus, for May and the EAP teachers, the issue of reporting (see 6.3b and 6.3c above) would diminish,

and the issue of their professional response profiled in its place. Figure 8.2 sets out examples of what such a response might involve.

Such an approach, however, is unlikely to be facilitated by the cycle of monitoring perceived by teachers as ‘a bureaucratic one-way street’ (see 6.4a above). Figure 8.3 sets out a twin-cycle approach which addresses the perspectives of teachers, programme development and quality assurance.

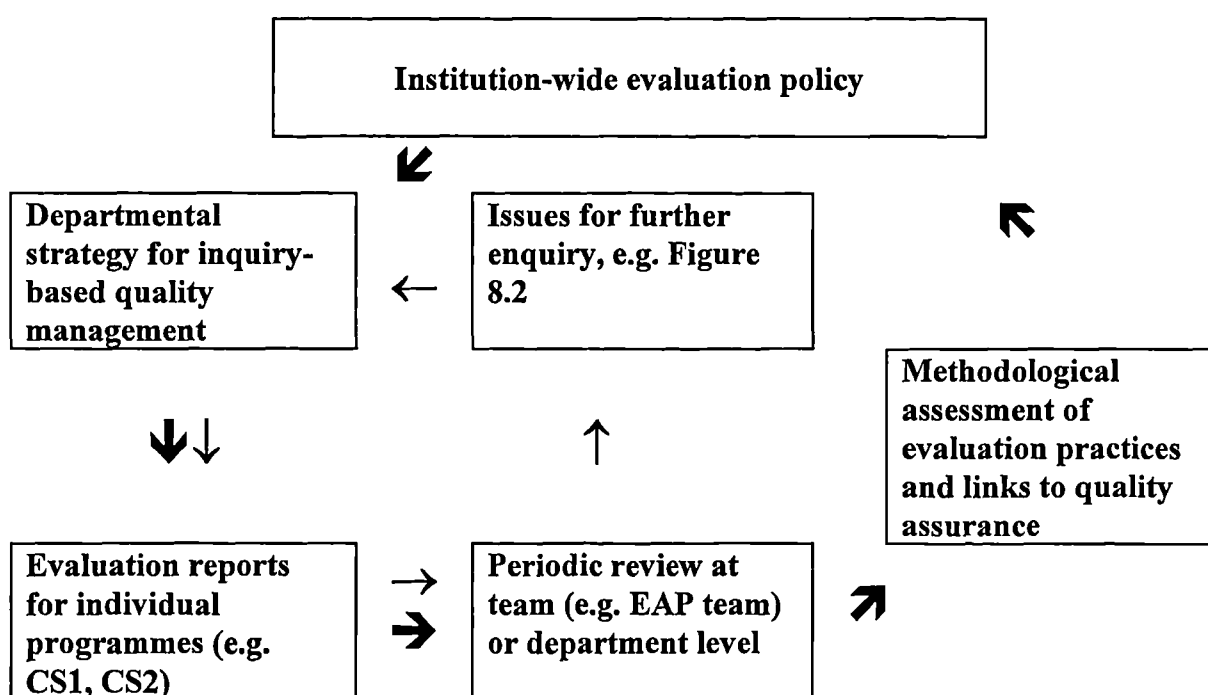


Fig. 8.3 Summary of evaluation policy development

Key assumptions underpinning this approach are: i) quality is a feature of the programme experience - the process rather than the design (the formal specification - see 5.2a above) or the outcomes (the examination board matrix), and ii) the primary contributors to programme quality are teachers and teaching teams. Thus, Anna's own professional contribution to CS2 (see 5.2 and 5.4 above) would be recognised and supported. The short cycle (represented by → in Figure 8.3) would focus on generating studies such as those outlined in Figure 8.2 above. The manager role

would be to secure resources and the involvement of other units of the institutional structure. The longer cycle (represented by ➔ in Figure 8.3) represents an on-going review of the evaluation and quality management framework. The key function here is methodological assessment - a process where the institution-wide framework is informed by policy and practice within the shorter cycle for departmental quality management.

The task here is not to set out detailed terms of reference for the fora and individuals which would be involved in these review cycles. Rather it is to show how the basic building blocks for quality management in the institutional context of CS1 and CS2: the grounding of quality in the experience of stakeholders; the use of evaluation for programme quality enhancement; the democratic approach to evaluation; and the concern for the perspective of external stakeholders (see 6.4a above and Appendix 6), might work differently to alleviate the problems with evaluation discussed in this thesis (see 5.6 above for a summary of these in the context of CS2). I suggest three strategies towards this end:

- i) The involvement of a teacher or manager from elsewhere in the institution to provide externality - a critical perspective and view on practices in other departments - and to link the two cycles. This might on the one hand, involve commenting critically, in the case of CS2, on the appropriateness of the NGT procedure: issues such as the focus on the early weeks of the programme, and the experience of Asian students on other programmes. It might also involve informing on evaluation practices in other departments. On the other hand, this person would use these insights from the evaluation of EAP or English language programmes to inform the longer review cycle, the means by which the institution discharges its responsibility for quality assurance.
- ii) The meetings and discussions involved in implementing the shorter cycle might focus on oral reports and dialogues rather than written reports. Thus, teachers

would talk about their experience of a given programme, including what emerged from student feedback, and possible ways of taking issues forward would be minuted as potential *Issues for further inquiry*. May, for example, would not receive reports to 'sit on her file' (see 6.3b above), but would co-write them with teachers like Anna and Millie, and decide with the team and critical informant, which issues to raise in other fora.

- iii) The case for quality assurance vis-a-vis external bodies would be a shared responsibility. It would be grounded in quality enhancement - the kind of decisions Anna made in the context of CS2 - and linked to processes of teacher appraisal and professional development. Ultimately, the formal responsibility for quality assurance would be a rich account of quality enhancement in line with the mission of the institution and the professionalism of teachers.

To counter the reality of a teacher-manager divide, and resulting boundary management (see 6.2 above), this approach needs i) implementation and ongoing scrutiny, and ii) a sound ethical base. The latter should ensure that no individual, or group of individuals is harmed or disadvantaged by the process, no time or energy is taken up by bureaucratic tasks which add no value to, or take time from teaching, and no problem which may impact negatively on the programme experience of individual students is ignored. The key features of such an ethical framework are discussed next.

8.4 An ethical framework

In the absence of such a framework, teachers and students are likely to act to protect their particular interests. Thus, Anna does not assist students to articulate and resolve problems (see 3.5b above), or to probe issues raised (see 5.3a above), and is hesitant about reporting and sharing of findings with others (see 6.3b above). The students in many cases remain silent on issues which are important to them, but which if expressed might have a negative impact on the teacher or on the group as a whole, for example, Niki on the race issue (see 5.3c above), and Laure on the materials she does

not like (see 7.3 above). Inaction on the part of the institution might be seen in the lax approach to the implementation of the evaluation policy, particularly in coordination and dissemination of feedback to the different departments and individuals who try to develop a strategy for evaluation (see 6.3b above). Each of these stakeholders might be seen as having a short term interest served by inaction. Teachers manage their teaching more easily when they do not uncover problems which they cannot solve (students such as Rata and Sao with too low a level in English to benefit from the programme) or which might take a lot of time to deal with (students such as Laure with a low level of awareness of the nature of their EAP problems). Students such as Abe in CS1 (see 3.5b above) and Laure in CS2 (see 7.4 above), may fare better by not allowing themselves to be characterised by a particular problem. The institution may find it appropriate not to query in detail how the evaluation policy is operating at a time when there are few resources for either a study of the current policy, or the development of an alternative. The convincing case presented to an external stakeholder for the efficacy of the evaluation policy in this context (see 3.2 above), and the problematic aspects of it emerging from this research (see Figure 8.1 above) together show the benefits to the department of not focussing on problems. While such benefits may be real in terms of external stakeholder judgements, they can be shown to be illusory and short-term, and thus ultimately untenable as an approach to the management of educational programmes. One strategy for dealing with problems which are perpetuated through inaction is the establishment of an ethical framework for programme evaluation and management. The discussion in this study suggests three ways in which an ethical framework might be established.

First, in a system such as that outlined above, the nature of *quality* might be better defined. This might be done as a development of the institution-wide commitment to being *student-led*. What this means for units such as recruitment of students onto programmes, and examinations would be clarified, and the problems Anna experienced with Sao and Rata (see 5.3a above), and Mari, Sona and Sue (see 6.2a above) might not inevitably be left to the teacher to resolve.

Second, to establish a culture of evaluation, *trust* is needed. The analysis of the wider context (see 6.4 above) suggests that this cannot be developed in one team or institution without regard for sector-wide construction of interests. It can, however, be initiated by open engagement with the responsibilities and possibilities of teachers and other stakeholders, by discussion in administrative and management fora of the role of evaluation, of following up evaluation findings, and of investment in the development of evaluation functions and practices. Thus, it would enhance the possibility of Anna seeing evaluation as a professional discourse, supporting her engagement with students' needs, rather than an institutional discourse which simplistically links satisfactory evaluation findings to *quality*, and problematic findings to *complaints* and possibly, *unsatisfactory teaching*. Teachers as a community might accept on the one hand, that evaluation findings are not used for informal appraisal purposes, and on the other, that there is a potential link between the practice of evaluation and the formal appraisal and staff development systems.

Third, the policy on evaluation needs to engage with the complex issue of programme ownership. For teaching, this means reconciling the views of teaching as *delivery* of a programme designed and owned by the institution, and teaching as constructed by the teacher, an *interpretation* of the specified programme. For management, that task is definition of processes which are enhanced by data from the student experience and from teachers, and the development of strategies which support a partnership towards a common mission. These two aspects are looked at in the next section.

8.5 Evaluation and teaching

Teaching involves evaluation in two ways. First, teachers interpret planned programmes and select materials and activities as an integral part of their teaching (see 5.2 above). Second, they formally evaluate the programme for purposes other than the construction of their pedagogy - as described in 3.5 and 5.3 above. The first type of evaluation is one which is recognised as a teaching skill, both in ELT and in

other educational contexts. The teacher's contribution to the learning programme through these evaluations and decisions is increasingly recognised as central to good teaching. The second type of evaluation is often seen as separate from the first; a service to programme and institutional management outside the classroom. The programmes described in this study show how the situation is more complex: there is both engagement with, and resistance to, the notion of evaluation as a pedagogic strategy (see 5.3a above). The engagement involves persuasion, for example, the use of the self-access learning resources in CS1 (see 3.5b above); awareness-raising, for example Anna's points about reading skills and text length in CS2 (see 5.3a above); and changes to pedagogy, for example the vocabulary teaching strategy in CS2 (see 5.5a above). The resistance is evident in what be described as a negative frame, in Anna's lack of commitment (see 5.3b above) and the overall perspective of Millie that the evaluation does not change the thrust of programme development (see 3.5c above). Resistance in a positive frame, i.e. one that links student feedback to coherent professional development, might be Anna's response to the vocabulary issue (see 5.5b above).

There are issues here for teacher education and policy development on teaching-related issues at institutional level. Teacher education might place greater emphasis on the role of context in representations of the teaching task. This is increasingly common in mainstream teacher education in Britain, where the particular characteristics and emphases of the National Curriculum are a key focus in teacher education programmes approved by the Teacher Training Agency. Thus, contextual features and policy frameworks are prioritised over context-free instructional techniques. In ELT there is a tradition of generalist training, with a focus on classroom-based, context-free instructional strategies. Teachers learn through practice of the significance of factors beyond the classroom. The response is to see these factors as intrusive, and strive to keep them at bay, (see 6.3 and 6.4 above where, in the case of Anna and May, this view of teaching seems to merge with the sector-wide teacher - manager divide to blunt the impact of the evaluation process). This view of

teaching is represented in the higher education context of CS1 and CS2 by the notion of teachers *delivering* programmes (see 8.4 above and 8.6 below). This study suggests two ways of improving this situation. First, continuing education for teachers needs to engage with the institutional policy frameworks which impact on teaching. Second, institutional policies for evaluation and other curricular processes might develop from a better informed view of teaching and the needs of teachers. The notion of 'student-led' might be replaced by a stakeholder approach which worked with the perspectives of all stakeholders. This latter point is developed further in the next section.

8.6 Programme evaluation: towards an integrated management strategy

There are two goals for the management of programmes within the institution. The first is to ensure that the programme is implemented - for example, the planning involved in staffing, timetabling, recruitment and placement, assessment, and provision of necessary resources such as materials and computers. The second is to ensure that the process of the programme constitutes a high quality learning experience for the students. The analysis of CS1 and CS2 above show how manager - teacher boundaries are managed in relation to both goals. In each case the responsibility for the programme is constructed as being with the institution - the programme specification, the timetable, the language centre, the examination arrangements and the evaluation policy. The teachers then, are responsible for *delivery* - a term which resonates of the technical end of service provision. The analysis of CS2, however, permits a different reading of the programme: the key features of the programme are actually Anna's own personal and professional contribution - the materials; the learning focus; management of discussions and presentation in the classroom. These are the aspects of the programme which the NGT generates feedback on. The students comment frankly and directly, as if they were somehow unconnected with Anna, and her role was simply to note their views. Anna is, in turn, silent, defensive, and finally accommodating of the changes proposed (see 5.3a above). When Sue provides positive feedback on Anna (see 5.3d above) it is graciously received, but not noted. It is as if Anna feels this feedback appraises her as

a programme input, like the timetable. The tension here reflects a wider conflict in higher education: the legislated accountability of the corporatised institution versus the notion of the autonomous teacher, whose authority derives from both subject expertise and professional duty (see 6.4b above).

An integrated programme management strategy in this context is one which does not query teacher ownership of programmes. Rather, it supports and develops this notion, such that the teacher's role and contribution are the basis for quality enhancement and quality assurance. To achieve this all units of the institution need to engage with the pedagogic, actively seeking to understand the issues which affect teaching and learning. The role for managers set out in 8.3 above illustrates this. Teachers, in turn need to quality manage their programmes. This means sharing their particular conception of a quality programme, the strategies they engage to achieve this, and the problems that arise. It means not only answering the question *Are students satisfied with the programme?* but also *Should students be satisfied with the programme?*

The four sets of issues discussed in 8.2 - 8.5 represent practical and conceptual building blocks towards this programme management culture. In relation to the practical, a more detailed fitting to context is needed than space permits. The conceptual dimension relates to the purpose: what teachers and managers want from evaluation. Young (1992) interpreting Habermas for the purpose of understanding classroom discourses, sets out 'three forms of knowledge-producing purposes or interests' which can guide the integration of discourses of evaluation on the one hand, and the different functions of evaluation examined in this study on the other. These are set out in Figure 8.3, with questions which critique evaluation from the perspectives of the main protagonists: teachers and managers. The interests in the first column represent ideological positions, which parallel the paradigm debate in evaluation (see 4.2 above), and which determine an interpretation of the evaluation policy. In the context of this study, Anna might see herself, as an EAP teacher (see 5.4 above) in the third row, and the institution in the first row. Managers responsible for

developing a ‘student-led’ institution might see the opposite: self-serving teachers obstructing the policies which might emancipate students. Answering the questions in the second and third columns might, as an ongoing constructive critique of evaluation in the context, might ultimately offer the possibility of synthesis.

Knowledge-producing interests (Young discussing Habermas)	Evaluation from a management perspective	Evaluation from a teacher’s perspective
An interest in control with a positivist self-understanding of the sciences and with the world of work	To what extent is evaluation and quality management compromising ownership of programmes by participants through forms of control?	To what extent is evaluation practice threatening teacher’s control, or serving to maintain teacher’s control of the programme?
An interest in understanding, associated with the hermeneutic sciences and cultural processes	To what extent is the organisation learning through the evaluation and quality management process?	To what extent is teacher-focussed programme development supported by evaluation practice?
An interest in emancipation, associated with the critical sciences and progressive social evolution. (after Young 1992:30)	To what extent is evaluation facilitating all participants realise their teaching and learning potential?	To what extent is evaluation assisting each teacher achieve what she wants, and assisting each student to learn more effectively?

Fig 8.4 Framework for evaluation development

8.7 Limitations of this study

This research study has undertaken a multi-dimensional examination of an already broad and complex topic and a number of limitations arise from this breadth of canvas. A holistic study of an EAP programme is difficult to achieve, and I am aware of having left some aspects undocumented. First, there are the ‘remote managers’ whom I did not interview, but who, as I analysed the data seemed to become more significant. Thus, they remain faceless, and possibly ‘sinister’, an unfair contrast with

the teachers and students. Second, while the data sets used may have depth, they are narrow, providing a less than ideal base for developing and confirming grounded hypotheses which relate to the wider contexts of the department, institution and higher education sector. I am mindful of the danger of over-interpreting from these data sets, and of the possibility that such awareness may not be adequate protection. Third, I am aware of a conceptual drift, as my analysis moved from the relatively bounded context of ELT into the more amorphous context of higher education. I could plead *grounded analysis*, but that does not quite bridge the focus of the literature chapter (Chapter 2) and the orientation of the later chapters which discuss the findings and their implications. Fourth, I make some recommendations for evaluation practice. While I accept that a study of this scope and methodological orientation cannot generalise, it can suggest further inquiry. In a field with a wealth of theorising, principles and guidelines for practice, more empirical studies of practice are required, and a pre-requisite for these is different forms of practice.

8.8 Concluding remarks

This chapter begins with the notions of connectedness and dislocation. These features of ELT programmes have been for me both experienced phenomena and intellectual curiosities. The curricular aspect of evaluation has proved an effective prism, providing views of often unseen dimensions of programmes. One of these dimensions is the notion of policy, a hazy area, neither theory nor practice. These foci, evaluation and policy, have in the EAP programmes studied here, illuminated diverse programme elements: the nature of programme quality, the different programme constructs which stakeholders hold, the power plays and survival strategies. The illumination is brief, and may not reach all the corners and crevices. Hopefully, it is enough to stimulate evaluation practice and related inquiry, to guide others wishing to research ELT and higher education programmes, and to encourage those teachers who feel quality management of programmes is an alien discourse. It is probably, as in Anna's case, what they do best!

Endnotes

Chapter 2

1. This phenomenon, however, is not limited to overseas evaluations: Mitchell (1992), discussing the Bilingual Education Project evaluation in Scotland, provides a rare account of how the evaluation focus is determined by political factors rather than the evaluator's consideration of what the study requires.
2. There is some evidence that the role of such general principles is changing: Holliday (1994; 1998) sets out a position where the role of student-student interaction in the English language classroom is, or should be determined by a range of contextual factors.
3. Legutke and Thomas (1991) locate their approach in a humanistic tradition of language education, which has paralleled the psycholinguistic, SLA-led approach of Richards and Rodgers over recent decades, rather than emerged as a reaction to it.
4. The programmes studied in Chapters 3 and 5 below represent such approaches, though in implementation, they become examples of individual teachers' notions of programme.
5. This view that experimental design is appropriate for research projects but not for evaluation is seen as problematic in two ways: first it assumes an epistemological difference between evaluation and research, an issue which is explored further in Chapter 4 below. Second, it reflects what Parlett and Hamilton label the agricultural-botany paradigm of understanding the effects on different treatments. This assumes that the features of an educational programme are as stable and comparable as treatments such as doses of nitrogen or hours of sunlight to crops growing in experimental conditions.
6. The Harris view here represents what is sometimes labelled evaluation-research - (Rea-Dickins 1994; Lynch 1996) - evaluations which have a strong theoretical or explanatory function (See Section 2.2 - Phase 3 evaluations). In terms of domains of evaluation practice, this category tends to prevail in the context of publicly-funded language education and social programmes.

Chapter 3

1. Formerly a Polytechnic, the institution became a university in 1992. This heralded a period of rapid expansion, with revision of curriculum management systems in order to deal with increased student numbers, and enhance learning and the quality of the student experience.
2. The British Association of State English Language Teaching (BASELT formerly BASCELT) is a body which audits and approves institutional capacity for English language programmes. The Higher Education Funding Council for England

(HEFCE) is a body which provides funding for programmes included within the undergraduate curriculum - such as the EAP programmes studied in this thesis - and which carries out periodic inspections of subject areas.

3. The term 'programme' is used within the institution to represent all the courses or modules taken by a student. In this thesis the term programme is used to represent the entity which is the focus of the evaluation, the English for Academic Purposes module.
4. In order to preserve anonymity, the names and the gender of teachers have been randomly allocated, and thus do not represent actual identity or gender. The names of students have been changed, but gender and nationality are retained, as these are considered significant factors in role development and participation in the programme.
5. The notion of culture here, and indeed throughout this thesis, is the set of values, practices and thought processes which members of a community share. The two professional communities here are university teachers on the one hand, and university managers on the other. This notion of culture is similar to the 'small culture' paradigm of Holliday (1999).
6. This term is used to refer to the activity within the programme. It does not suggest an analysis of language use such as is common within the English for Special Purposes (ESP) field.
7. The data referencing convention followed in this study identifies data type (here classroom notes) the programme case study (here case study 1) and the week (here Week 1).
8. See Section 4.8 for a discussion of how, following Lincoln and Guba (1985), reliability is considered as a dimension of validity in this thesis.

Chapter 4

1. The use of "case" for research focus is revealing: it suggests that the research is being carried out in a context where that context is important. The guiding principle in positivistic research is to inform on processes universally, to seek out laws which apply to all contexts (see Section 4.4).
2. As I identified the students after the first session I provided each with a pseudonym, such that I could discuss them as real people rather than as an initial or number in the course of the study. To maintain this anonymising I used the real name, or abbreviation thereof, in observation notes and interview transcripts, in versions for teachers and students to check, and switched to the pseudonym as I finally added texts to the research data sets.

3. *Note on transcription*

This transcript represents the language of the interview rendered into Standard English. While all the actual words are included, hesitations and minor grammatical mistakes and infelicities are changed, and the phonological features of the language are not transcribed. There are three reasons for this: no detailed discourse or conversational analysis which might require such features is anticipated, it is easier for the students themselves to read, augment and confirm a text which seems correct, and I did not want to represent the students through inaccurate English - which they might see as a negative representation. A common tendency with student member checks of transcripts in a previous research project was to correct minor points relating to articles, etc. which I left unaltered, and to improve on word choice and grammar.

Chapter 5

1. *A note on terminology*

The term 'event' is selected to represent the evaluation described here. It includes the instrument (the three questions), its implementation and the discussion of the data emerging. Where the NGT is referred to more generally, as a construct, it is termed a 'procedure'.

2. Niki's assertion that her experiences are inexpressible might seem to conflict with the clarity with which she articulates them in the interview. She feels that she can say things in an interview context with me (she was a student in my class on a summer programme the previous year):

Niki: Yes, I think so, but this [the interview context] is very different to evaluation in the classroom, because there are other sort of things, it's quite difficult for my own personal feelings ...(Interview with Niki CS2:7)

3. Jin (1992) quoted in Cortazzi and Jin (1996) propose a Cultural Synergy approach to integrating Asian and Western cultural perspectives in such English programmes in British universities. In some respects, Anna's 'critical' approach to the programme (see 5.2 above) incorporates elements of this cultural awareness approach, but it does not seem to alleviate the issues which derive from perceived cultural differences by students.
4. Laure is referring her to the fact that it is my copy of the text which we are looking at in the interview, not her annotated and colour-coded copy. See Section 4.2b for a discussion of the interview strategy.
5. In the field of language testing, where Messick developed his framework, the consequences of a test result for an individual, for example, acceptance on a university programme, are often more clearly established than are the consequences for any one individual of an evaluation such as the NGT in CS2.

Chapter 6

1. Fullan is describing regulatory processes in schooling at primary and secondary level. Norris (1998) extends this characterisation of regulation of educational programmes to the British higher education context.
2. *Etic* and *emic* are used here in a metaphorical sense, characterising somewhat simplistically the key difference between the two approaches, in a manner which corresponds broadly to the approach for data analysis developed for this study.

Chapter 7

1. While Laure's participation in these tasks is not strategic in the Flavell/Wenden language learning sense, it does represent a strategic perspective in relation to her social construction of the programme (see 7.2 above). The activity provided an opportunity for energetic, collaborative problem-solving, with me as an informant. Laure's account of her participation in the vocabulary activity 30/4/2 (see Section 5.5.a), similarly shows her maximising the interactional implications of classroom activities.

Bibliography

- Ackers, J. 1997 Evaluating UK courses: the perspective of the overseas student. in D. McNamara & R. Harris (Eds.) pp. 107-200
- Adelman, C. 1993 Kurt Lewin and the origins of action research. *Educational Action Research* Vol 1 No 1 pp. 7-24
- Adelman, C, D. Jenkins, & S. Kemmis 1981 Rethinking case study: notes from the Second Cambridge Conference. in H. Simons (Ed.) pp. 45-61
- Alderson, J .C. 1992 Guidelines for the evaluation of language education. In J. C. Alderson & A. Beretta (Eds.) pp. 274-304
- Alderson, J.C. & A. Beretta (Eds.) *Evaluating Second Language Education*. Cambridge: Cambridge University Press
- Alderson, J.C. & M. Scott 1992 Insiders, outsiders and participatory evaluation. in J. C. Alderson & A. Beretta (Eds.) 1992 pp. 25-58
- Allwright, D. & K. Bailey 1991 *Focus on the Language Classroom*. Cambridge: Cambridge University Press
- Bailey, K.M. 1996 The best laid plans. in K. M. Bailey & D. Nunan (Eds.) pp. 15-40
- Bailey, K.M. & D. Nunan (Eds.) 1996 *Voices from the Language Classroom*. Cambridge: Cambridge University Press
- Barnes, D. 1976 *From Communication to Curriculum*. Harmondsworth: Penguin
- Barnett, R. 1994 *Limits of Competence*. Milton Keynes: Open University Press / The Society for Research into Higher Education
- Becher, T. 1989 *Academic Tribes and Territories*. Milton Keynes: Open University Press / The Society for Research into Higher Education
- Beretta, A. 1986a Towards a methodology of ESL programme evaluation. *TESOL Quarterly* Vol 20 No 1 pp. 144-155
- Beretta, A. 1986b Programme-fair language teaching programme evaluation. *TESOL Quarterly* Vol 20 No 3 pp. 431-444
- Beretta, A. 1986c A case for field experimentation in programme evaluation. *Language Learning* Vol 36 No 3 pp. 295-309
- Beretta, A. 1987 Evaluation of a language-teaching project in South India. Ph.D. Thesis, University of Edinburgh

- Beretta, A. 1989a Attention to form or meaning?: Error treatment in the Bangalore Project. *TESOL Quarterly* Vol 23 No 2 pp. 283-303
- Beretta 1989b Who should evaluate L2 Programmes? In C. J. Brumfit & R. Mitchell (Eds.) pp. 155-160
- Beretta, A. 1990b Implementation of the Bangalore Project. *Applied Linguistics* Vol 11 No 1 pp. 1-15
- Beretta, A. 1992a Evaluation of language education: an overview. In J. C. Alderson and A. Beretta (Eds.) pp. 5-24
- Beretta, A. 1992b What can be learnt from the Bangalore evaluation? In J. C. Alderson and A. Beretta (Eds.) pp. 250-271
- Beretta, A 1992c Editor's postscript to Alderson and Scott 1992. in Alderson and Beretta (Eds.) pp. 58-61
- Beretta, A. & A. Davies 1985 Evaluation of the Bangalore Project. *English Language Teaching Journal* Vol 39 No 2 pp. 121-127
- Block, D. 1996 A window on the classroom: classroom events viewed from different angles. in Bailey & Nunan (Eds.) pp. 168-194
- Block, D. 1998 Tale of a language learner. *Language Teaching Research* Vol 2 No 2 pp. 148-176
- Bloom, D & J. Green 1997 Ethnography and ethnographer of and in education: a situated perspective. in J. Flood, S. Heath & D. Lapp 1997
- Blue, G. & P. Grundy 1996 Team evaluation of language teaching and language courses. *English Language Teaching Journal* Vol 50 No 3 pp. 244-253
- Bochner, S., B. McLeod & A. Lin (1997) Friendship patterns in overseas students: a functional model. *International Journal of Psychology* Vol 12 pp. 277-299.
- Bottery, M. 1992 *The Ethics of Educational Management*. London: Cassell
- Bowers, R. 1983 Project planning and performance. *ELT Docs* 116 pp. 90-120
- Breen, M. 1985 The social context of language learning - a neglected situation? *Studies in Second Language Acquisition* Vol 7 No 2 pp. 135-158
- Breen, M. & C. Candlin 1980 The essentials of a communicative curriculum in language teaching. *Applied Linguistics* Vol 1 No 1 pp. 89-112
- Britten, D. & J. O'Dwyer 1995 Self-evaluation in in-service teacher training. in Rea-Dickins & Lwaitama 1995 pp. 87-106

- Brown, J.D., 1989 Language programme evaluation: a synthesis of existing possibilities. In R. K. Johnson (Ed.) 1989 pp. 242-261
- Brown, S. & P. Race 1994 Assess your own Teaching Quality. London: Kogan Page
- Brumfit, C.J. 1984a Communicative Methodology in Language Teaching - The role of accuracy and fluency. Cambridge: Cambridge University Press
- Brumfit, C.J. 1984b The Bangalore Procedural Syllabus. English Language Teaching Journal Vol 38 No 4 pp. 233-241
- Brumfit, C.J. & R. Mitchell 1989 Research in the Language Classroom. ELT Docs 133. Modern English Publications
- Burrell, G. & G. Morgan 1979 Sociological Paradigms and Organisational Analysis. London: Heinemann
- Burstall, C., M. Jamieson, S. Cohen, & M. Hargreaves 1974 Primary French in the Balance. Slough: National Foundation for Educational Research
- Bush, T. 1995 Theories of Educational Management. London: Paul Chapman
- Calderhead, J. 1988 Teachers' Classroom Decision Making. London: Cassell
- Canagarajah, A.S. 1996 Review of A. Holliday 1994 Appropriate Methodology and Social Context. English Language Teaching Journal Vol 50 No 1 pp. 80-82
- Canale, M. and M. Swain 1980 Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics Vol 1 No 1: 1-47
- Candlin, C. & D. F. Murphy 1987 Language Learning Tasks. Oxford: Pergamon
- Carr, W. & S. Kemmis 1986 Becoming Critical. Lewes: Falmer Press
- Celani, M. M. A., J. L. Holmes, R. C. G. Ramos, M. R. Scott 1988 The Brazilian ESP Project: An Evaluation. Sao Paulo: Editora de PUC-SP
- Chambers, F. 1995 Cyclic innovation in ELT: implications for implementation. in J. O'Dwyer (Ed.) pp. 67-76
- Chater, M. F. T. 1998 The Gods of Mismanagement: a mythological exploration of our management predicament. in School Leadership & Management Vol 18 No 2: 231-238
- Chomsky, N. 1959 Review of Skinner (1957). Language, Vol 35 pp. 26-28
- Clandinin, D. J. 1986 Classroom Practice: Teacher Images in Action. Philadelphia: Falmer Press

- Clark, R., A. Fortune, R. Kiely & J. O'Regan 1996 Integrated British Studies and EFL: students' perceptions. IATEFL Literature and Cultural studies SIG Newsletter Issue 13 pp. 21-33
- Clark, R. & R. Ivanic 1991 Consciousness-raising about the writing process. in C. James & P. Garnett (Eds.) pp. 168-185
- Cohen, L. & L. Manion 1994 Research Methods in Education (Fourth Edition). London: Routledge
- Connelly, F. M. & D. J. Clandinin 1988 Teachers as Curriculum Planners: Narratives of Experience. New York: Teachers College Press
- Cooper, L. 1995 The relationship between training, self-evaluation and external evaluation for NGOs in Central and Eastern Europe. Paper presented at the International Evaluation Conference, Vancouver 1995
- Cortazzi M. & L. Jin 1996 English Teaching and Learning in China. Language Teaching Vol 29 No 2 pp. 61-80
- Crookes, G. 1993 Action Research for second language teachers: going beyond teacher research. Applied Linguistics Vol 14 No 2 pp. 130-144
- Crossley, M. & J.A. Bennett 1995 Planning for case study evaluation in Belize, Central America. In M. Crossley & G. Vulliamy (Eds.) pp. 138-157
- Crossley, M. & G. Vulliamy (Eds.) 1995 Qualitative Research in developing Countries: educational Perspectives. New York: Garland
- Cuthbert, R. 1996 Working in Higher Education. Milton Keynes: Open University Press / The Society for Research into Higher Education
- Day, R. (Ed.) 1986 Talking to Learn. Rowley, Mass.: Newbury House
- Delbecq, A. 1975 Group Techniques for Programme Planning: A Guide to Nominal Group and Delphi Processes. Glenview, Ill.: Scott Foreman & Co.
- Doloughan, F. 1999 English as a second thought: Letter to the Editor. Times Higher Education Supplement, June 25 1999
- Dubin, F. & R. Wong 1989 An ethnographic approach to inservice preparation: the Hungary file. in Richards & Nunan (Eds.): 282-291
- Duff, P. 1986 Another look at interlanguage talk: taking task to task. in R. Day (Ed.) 147-181
- Edge, J. & K. Richards 1998 May I see your warrant, please?: Justifying outcomes in qualitative research. Applied Linguistics Vol 19 No 3 pp. 334-356

- Edge, J. & K. Richards (Eds.) 1999 Proceedings of Teachers Develop Teachers' Research Conference (TDTR4) Leuven Belgium. (CD - ROM)
- Edwards, A. D. & Westgate, D. P. G. 1987 Investigating Classroom Talk. Lewes: Falmer Press
- Ellen, R. 1984 Ethnographic Research: a Guide to General Conduct. London: Academic Press
- Elliot, J. 1991 Action Research for Educational Change. Milton Keynes: Open University Press
- Elliot, J. (Ed.) 1993 Reconstructing Teacher Education. Lewes: Falmer Press
- Fetterman, D. M. 1988 Qualitative Approaches in Evaluation in Education: The Silent Revolution. New York: Praeger
- FitzGibbon, C. T. 1996 Monitoring Education: Indicators, Quality and Effectiveness. London: Cassell
- Flavell, J. H. 1979 Metacognition and cognitive monitoring: a new era of cognitive developmental inquiry. *American Psychologist* Vol 34 pp. 906-911
- Flood, J., S. Heath & D. Lapp 1997 (Eds.) A Handbook of Research on Teaching Literacy through the Communicative Visual Arts. New York: Macmillan
- Fullan, M. 1991 The New Meaning of Educational Change. London: Cassell
- Furedi, F. 1999 Mind your English Language. *Times Higher Education Supplement*. 11 June 1999 p.18
- Furnham, A. 1997 The experience of being an overseas student. in D. McNamara & R. Harris (Eds.) pp. 13-30
- Gardner, R.C., P.D. MacIntyre & L.M. Lysynchuk 1990 The affective dimension in second language programmes evaluation. in *Language, Culture and Curriculum* Vol 3 No 1:39-64
- Gibbs, G. 1992 Developing Teaching: Teaching More Students. Oxford: Oxford Centre for Staff Development
- Gibson, R. 1985 Critical times for action research. *Cambridge Journal of Education* Vol 15: 59-64
- Giddens, A. 1989 Sociology. Oxford: Polity Press
- Gitlin, A. & J. Smyth 1989 Teacher Evaluation: Educative Alternatives. Lewes: Falmer Press

Glass, G. V. & N. Worthen 1971 Evaluation and research: similarities and differences. Curriculum Theory Network. Fall pp. 149-165

Goetz, J. & M. LeCompte 1981 Ethnography and Qualitative Design in Educational Research. San Diego: Academic Press

Goodlad, S. 1995 The Quest for Quality. Milton Keynes: Open University Press / The Society for Research into Higher Education

Graves, K. 1996 Teachers as Course Developers. Cambridge: Cambridge University Press

Green, D. 1994 (Ed.) What is Quality in Higher Education? Milton Keynes: Open University Press / The Society for Research into Higher Education

Green, D., C. Brannigan, P. Mazelan & L. Giles 1994 Measuring student satisfaction: a method of improving the quality of the student's experience? in S. Haselgrove (Ed.) pp. 99-107

Green, J. & D. Bloome 1997 Ethnography and ethnographers of and in education; A situated perspective. in J. Flood, S. B. Heath & D. Lapp (Eds.) 1997

Greenwood, J. 1985 Bangalore revisited: a reluctant complaint. English Language Teaching Journal Vol 39 No 4 pp. 268-273

Guba, E. & Y. Lincoln 1989 Fourth Generation Evaluation. Newbury Park: Sage

Habermas, J. 1968 Knowledge and Human Interests (translated by J. Shapiro 1972). London: Heinemann Educational

Habermas, J. 1987 The Theory of Communicative Action Vol 1. (translated by T. McCarthy 1989) Cambridge: Polity

Hadfield, J. 1997 Correspondence. English Language Teaching Journal Vol 51 No 3 pp. 310-312

Hammersley, M. 1992 What's Wrong with Ethnography? London: Routledge

Hammersley, M. 1993 On the teacher as researcher. Educational Action Research Vol 1 No 3 pp. 425-445

Hammersley, M. & P. Atkinson 1983 Ethnography: Principles in Practice. Tavistock Institute of Human Relations

Harris, J. 1990 The second language programme-evaluation literature: accommodating experimental and multi-faceted approaches. Language, Culture and Curriculum Vol 3, No 1 pp. 83-92

- Haselgrove, S. (Ed.) 1994 *The Student Experience*. Milton Keynes: Open University Press / The Society for Research into Higher Education
- Henry, 1994 An ethical perspective. in S. Haselgrove (Ed.) pp. 108-116
- Hollander, S. 1965 *The Sources of Increased Efficiency: a Study of Du Pont Rayon Plants*. Boston: MIT Press
- Holliday, A. 1994 *Appropriate Methodology and Social Context*. Cambridge: Cambridge University Press
- Holliday, A. 1995 Evaluation as cultural negotiation. in Kiely *et al* (Eds.) 1995 pp. 6-11
- Holliday, A. 1997 Six lessons: cultural continuity in communicative language teaching. *Language Teaching Research* Vol 1 No 3 pp. 212-238
- Holliday, A. 1999 Small cultures. *Applied Linguistics* Vol 20 No 2 pp. 237-264
- Hopkins, D. 1985 *A Teacher's Guide to Classroom Research*. Milton Keynes: Open University Press
- Hopkins, D. 1989 *Evaluation for School Development*. Milton Keynes: Open University Press
- Hornberger, N. 1994 Alternatives in TESOL Research: Ethnography. *TESOL Quarterly* Vol 26 No 4 pp. 688-690
- House, E. R. (Ed.) 1986 *New Directions in Educational Evaluation*. Lewes: Falmer Press
- Ivanic, R. & R. Clark 1997 *The Politics of Writing*. London: Routledge
- James, C. & P. Garrett (Eds.) 1991 *Language Awareness in the Classroom*. Harlow: Longman
- Johnes, J. & J. Taylor 1990 *Performance Indicators in Higher Education*. Milton Keynes: Open University Press / The Society for Research into Higher Education
- Johnson, K. E. 1995 *Understanding Communication in Second Language Classrooms*. Cambridge: Cambridge University Press
- Johnson, R.K. (Ed.) 1989 *The Second Language Curriculum*. Cambridge: Cambridge University Press
- Karavas-Doukas, E. 1996 Using attitude scales to investigate teachers' attitudes to the communicative approach. *English Language Teaching Journal* Vol 50, No 3 pp. 187-199

- Karavas-Doukas, E. 1998 Evaluating the implementation of educational innovations: lessons from the past. in Rea-Dickins, P. and K. Germaine (Eds.) 1998 pp. 25-50
- Kemmis, S. & R. McTaggart 1981 *The Action Research Planner*. Victoria: Deakin University Press
- Kennedy, C. 1988 Evaluation of the management of change in ELT projects. *Applied Linguistics* Vol 9 No 4 pp. 329-342
- Kennedy, J. 1999 Using mazes in teacher education. *English Language Teacher Education* Vol 53 No 2 pp. 107-114
- Kiely, R. 1996 Professional development for teacher trainers: a materials writing approach. *English Language Teaching Journal* Vol 50 No 1: 59-66
- Kiely, R. 1998 Programme evaluation by teachers: issues of policy and practice. in Rea-Dickins and Germaine 1998 pp. 78-104
- Kiely, R. 1999a Evaluation by teachers. *IATEFL Issues* 147 pp. 14-16
- Kiely, R. 1999b Taking stock: mid-course evaluation. *IATEFL Testing, Evaluation and Assessment SIG Newsletter* pp. 6-9
- Kiely, R. 1999c Classroom evaluation: a platform for teacher development. in J. Edge & K. Richards (Eds.) 1999 *Proceedings of Teachers Develop Teachers' Research Conference (TDTR4)* Leuven Belgium. CD - ROM
- Kiely, R., D. F. Murphy, P. Rea-Dickins & M. I. Reid (Eds.) 1995 *Evaluation in Planning and Managing Language Education Projects - Papers of the Second PRODESS Colloquium*. Manchester: The British Council
- Kiely, R., D. F. Murphy, P. Rea-Dickins & M. I. Reid 1995 *PRODESS Evaluation Guidelines*. Manchester: The British Council
- Kinnell, M. (Ed.) 1990 *The Learning Experiences of Overseas Students*. Buckingham: The Society for Research into Higher Education / The Open University Press
- Krashen, S. D. 1982 *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press
- Krashen, S. D. & T. D. Terrell 1983 *The Natural Approach: Language Acquisition in the Classroom*. San Francisco: Alemany Press
- Krueger, R. 1994 *Focus Groups: A Practical Guide*. London: Sage
- Kushner, S. 1995 *Democracy in Flight: an essay on case study, programme evaluation and political education*. Presented at First Annual Conference of the UK Evaluation Society

- Kushner, S. 1996 The limits of constructivism. *Evaluation* Vol 2 No 2 pp. 189-200
- Lawrence, L. 1990 Language in education: an evaluation of the teaching of structure in Zambian grade 8 classes. Unpublished PhD Thesis Lusaka: University of Zambia
- Lawrence, L. 1995 Using evaluation to improve teacher education programmes. in Rea-Dickins, P. & A. F. Lwaitama (Eds.) 1995
- Lea, M. R. & B. Street 1998 Student writing in Higher Education: an academic literacies approach. in *Studies in Higher Education* Vol 23 No 2 pp. 157-172
- Legutke, M. & H. Thomas 1991 Process and experience in the language classroom. Harlow: Longman
- Lemke, J. 1995 *Textual Politics: Discourse and Social Dynamics*. London: Taylor and Francis
- Lewin, K. 1946 Action research and minority problems. in G. W. Lewin (Ed.)
- Lewin, G. W. (Ed.) 1946 *Resolving Social Conflicts*. New York: Harper & Row
- Lewins, H. 1990 Living needs. in M. Kinnell (Ed.) pp. 82-106
- Li E. & G. James (Eds.) *Testing and Evaluation in Second Language Education*. Hong Kong: Hong Kong University of Science and Technology
- Lincoln, Y. & E. Guba 1985 *Naturalistic Inquiry*. Newbury Park: Sage
- Linn, R. N. (Ed.) 1989 *Educational Measurement*. New York: Macmillan
- Long, M. H. 1984 Process and product in ESL programme evaluation. *TESOL Quarterly* Vol 18 No 3 pp. 409-425
- Long, M. & P. A. Porter 1985 Groupwork, interlanguage talk and second language acquisition. *TESOL Quarterly* Vol 19 No 2 pp. 207-228
- Lynch, B. 1996 *Language Programme Evaluation*. Cambridge: Cambridge University Press
- Macdonald, B. 1976 Evaluation and the control of education. in Tawney, D. (Ed.)
- Macdonald, B. & R. Walker 1975 Case study and the social philosophy of educational research. *Cambridge Journal of Education* Vol 5 pp. 2-11
- Mackay, R. 1994a Undertaking ESL/EFL programme review for accountability and improvement. *English Language Teaching Journal* Vol 48 No 2 pp.142-149
- Mackay, R. 1994b Making evaluation work. *PRODESS Newsletter* No 6 pp. 5-6

- Mackay, R., S. Wellesley, & E. Bazergan 1995 Participatory evaluation. *English Language Teaching Journal* Vol 49 No 4 pp. 308-317
- Marsh, H. W. 1987 Students' evaluations of university teaching: Research findings, methodological issues and directions for future research. *International Journal of Educational Research* Vol 11 No 3 pp. 257-387
- McCarthy, M. & R. Carter 1995 Spoken Grammar: What is it and how can we teach it?. *English Language Teaching Journal* Vol 49 No 3 pp. 207-218
- McCarthy, M. & R. Carter 1996 Correspondence. *English Language Teaching Journal* Vol 50 No 4 pp. 369-371
- McDonagh, J. & C. Shaw 1994 *Materials and Methods in ELT*. Oxford: Backwell
- McHoul, A. 1978 The organisation of turns of formal talk in the classroom. *Language in Society* Vol 51 pp. 183-213
- McKay, V. & C. Treffgarne (Eds.) 1998 *Evaluating Impact*. Serial No 35, Department for International Development Education Papers
- McKinley, C. 1999 English as a second thought: Letter to the Editor. *Times Higher Education Supplement*, June 25 1999
- McNay, I. 1995 From the collegial academy to corporate enterprise; the changing cultures of universities. in T. Schuller (Ed.) pp.105-115
- McNiff, J. 1988 *Action Research: Principles and Practice*. Basingstoke: Macmillan Educational.
- Messick, S. 1989 Validity. in R. N. Linn (Ed.) pp. 13-103
- Messick, S. 1994 The interplay of evidence and consequences in the validation of performance instruments. *Educational Researcher* Vol 23 No 2 pp. 213-223
- Messick, S. 1996 Validity and washback in language testing. *Language Testing* Vol 13 pp. 241-256
- Miles, M. B. & A. M. Huberman 1984 *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills: Sage
- Miletto, H. & O. Raffner (Eds.) 1997 *Learning a Language Differently*. Strasbourg-Wacken: European Youth Centre
- Mitchell, R. 1989 Second language learning: Investigating the classroom context. *System* Vol 17, No 2 pp. 195-210
- Mitchell, R. 1990 Evaluation of second language teaching projects and programmes. *Language, Culture and Curriculum* Vol 3 No 1: 3-17

- Mitchell, R. 1992 The 'independent' evaluation of bilingual primary education. in Alderson & Beretta (Eds.) 1992 pp. 100-137
- Mitchell, R., B. Parkinson, R. Johnstone 1981 The Foreign Language Classroom: an observational study. Stirling Educational Monographs No 9. Department of Education, University of Stirling
- Munby, J. 1978 Communicative syllabus design. Cambridge: Cambridge University Press
- Murphy, D. F. 1996 The evaluator's apprentices: Learning to do evaluation. Evaluation Vol 2 No 3 pp. 321-328
- Murphy, D. F. & P. Rea-Dickins 1998 Identifying stakeholders. In V. McKay & C. Treffgarne (Eds.) pp. 89-90
- Norris, N. 1990 Understanding Educational Evaluation. London: Kogan Page / CARE, University of East Anglia
- Norris, N. 1998 Curriculum evaluation revisited. Cambridge Journal of Education Vol 28 No 2 pp. 207-220
- Nunan, D. 1989 Designing Tasks for the Communicative Classroom. Cambridge: Cambridge University Press
- Nunan, D. 1990 Action research in the language classroom. In Richards & D. Nunan (Eds.) pp. 62-81
- Oakley, A. 1999a Overheads. London: Flamingo
- Oakley, A. 1999b A Motley mirror image. in Times Higher Education Supplement, 26 March 1999: 20
- O'Dwyer, J. 1996 (Ed.) Quality Management and the Management of Change: Papers presented at an International Conference on Management in English Language Teaching. Ankara, Turkey: Bilkent University School of English Language
- Opacic, S. 1994 The student learning experience in the mid-1990s. in S. Haselgrove (Ed.) 1994:157-168
- O'Regan, J. 1997 Evaluation. in H. Miletto & O. Raffner (Eds.) pp. 65-88
- Oxford, R. 1994 Language Learning Strategies: What every teacher should know. Boston: Heinle & Heinle
- Parkinson, B., D. McIntyre, & R. Mitchell 1982 An independent evaluation of 'Tour de France'. Stirling Educational Monographs No 11. Department of Education, University of Stirling

- Parlett, M. & D. Hamilton 1972 Evaluation as illumination: a new approach o the study of innovatory programmes. Occasional Paper No 9. Centre for Research in the Educational Sciences, Edinburgh
- Parton, N. 1994 The nature of social work under conditions of (post) modernity. Social Work & Social Sciences Review Vol 5 No 2 pp. 93-112
- Patton, M. Q. 1982 Practical Evaluation. Thousand Oaks: Sage
- Patton, M. Q. 1997 Utilisation-Focussed Evaluation (3rd edition). Thousand Oaks: Sage
- Pawson, R. & N. Tilley 1997 Realistic Evaluation. London: Sage
- Peck, A. & D. Westgate 1994 Language Teaching in the Mirror. London: CILT
- Pennington, M. 1998 Designing language programme evaluations for specific purposes. in E. Li & G. James (Eds.) pp. 199-213
- Pennington, M. C. & A. L. Young 1989 Approaches to faculty evaluation for ESL. TESOL Quarterly Vol 23 No 4 pp. 619-646
- Pollard, A. & S. Tann 1987 Reflective Practice in the Primary School. Lewes: Falmer Press
- Popham, W. J. 1975 Educational Evaluation. Englewood Cliffs, New Jersey: Prentice Hall
- Prabhu, N. S. 1987 Second Language Pedagogy. Oxford: Oxford University Press
- Prodromou, L. 1996a Correspondence. English Language Teaching Journal Vol 50 No 1 pp. 88-89
- Prodromou, L. 1996b Correspondence .English Language Teaching Journal Vol 50 .No 4 pp. 371-373
- Race, P. 1994 The Open Learning Handbook. London: Kogan Page
- Rea-Dickins, P. 1994a Evaluation and English Language Teaching. Language Teaching Vol 27 No 2 pp. 71-92
- Rea-Dickins, P. 1994b Setting up and monitoring the evaluation process. PRODESS Newsletter 2 Manchester: The British Council / Thames Valley University
- Rea-Dickins, P. & K. Germaine 1992 Evaluation. Oxford: Oxford University Press

- Rea-Dickins, P & A.F. Lwaitama 1995 (Eds.) Evaluation for development in English language teaching: Review of ELT. Vol 3 No 3 Basingstoke: Modern English Publications/The British Council
- Rea-Dickins, P. & K. Germaine 1998 (Eds.) Managing Evaluation and Innovation in Language Teaching. Harlow: Longman
- Reyal, M. 1999 English as a second thought: Letter to the Editor. Times Higher Education Supplement, June 25 1999
- Richards, J. C. 1984 The secret life of methods. TESOL Quarterly Vol 18 No 1 pp. 7-23
- Richards, J. C. & T. S. Rogers 1986 Approaches and Methods in Language Teaching: a Description and Analysis. Cambridge: Cambridge University Press
- Richards, J. & C. Lockhart 1994 Reflective teaching in second language classrooms. Cambridge: Cambridge University Press
- Richards, J. C. & D. Nunan 1989 (Eds.) Second Language Teacher Education. Cambridge: Cambridge University Press
- Richardson, J. E. 1994 A British evaluation of the course evaluation questionnaire. Studies in Higher Education Vol 19 No 1 pp. 59-62
- Riddle, M. 1994 Undergraduate coursework assignments: matching quality to experience of outcome. Unpublished paper (used for teaching in CS2).
- Roberts, C. 1998 Language and cultural issues in innovation: the European dimension. In P. Rea-Dickins & K. Germaine (Eds.) pp. 51-77
- Roberts, C., E. Davies & T. Jupp 1992 Language and Discrimination. Harlow: Longman
- Roberts, C., C. Garnett, S. Kapoor & S. Sarangi 1992 Quality in Teaching and Learning. Sheffield: Training, Enterprise and Education Directorate, Department of Employment
- Rogers, E. M. 1983 The Diffusion of Innovations. London: Macmillan
- Rossner, R. & J. Strange 1995 Who knows what's best for language learners? Paper presented at 29th International Annual IATEFL Conference, York.
- Sanders, J. R. 1992 Evaluating School Programmes: An Educator's Guide. Newbury Park: Corwin Press
- Schön, D.A. 1983 The Reflective Practitioner. Aldershot: Arena

- Schuller, T. (Ed.) 1995 *The Changing University?* Milton Keynes: Open University Press / The Society for Research into Higher Education
- Senior, R. 1997a Transforming language classes into bonded groups. *English Language Teaching Journal* Vol 51 No 1 pp. 3-11
- Senior, R. 1997b Correspondence. *English Language Teaching Journal* Vol 51 No 4 pp. 310
- Shamin, F. 1996 In and out of the action zone: location as a feature of interaction in large ESL classes in Pakistan. In K. Bailey & D. Nunan (Eds.) pp. 123-144
- Simons, H. (Ed.) 1981 *Towards a Science of the Singular*. Norwich: Centre for Applied Research in Education, University of East Anglia
- Simpson, J. 1996 Of flux and functional fixedness: towards an educational role for EAP. *Research News - The Newsletter of the IATEFL Research Special Interest Group* No 8 pp. 15-21
- Skehan, P. 1996 A framework for the implementation of task-based instruction. *Applied Linguistics* Vol 17 No 1 pp. 38-62
- Skinner, B. F. 1957 *Verbal Behaviour*. New York: Appleton-Century Crofts
- Smith, L.M. 1978 An evolving logic of participant observation, educational ethnography and other case studies. *Review of Research in Education* Vol 6 pp. 316-377. Chicago: Peacock Press
- Somekh, B. 1989 The role of action research in collaborative enquiry. in B. Somekh, J. Powney, & C. Burge (Eds.)
- Somekh, B., J. Powney, & C. Burge (Eds.) *Collaborative Enquiry and School Improvement*. Norwich: Classroom Action Research Network, University of East Anglia
- Stake, R. 1995 *The art of case study research*. Thousand Oaks: Sage
- Stake, R. 1996 For all programme evaluators, there's the criterion problem. (Address to the Annual Conference UKES 1995). *Evaluation* Vol 2 No 1 pp. 93-101
- Stenhouse, L. 1975 *An introduction to curriculum research and development*. London: Heinemann
- Stern, H. H. 1983 *Fundamental Principles in Language Teaching*. Oxford: Oxford University Press
- Stoller, F. A. 1994 The diffusion of innovation in intensive ELT programmes. *Applied Linguistics* Vol 15 No 3 pp. 300-327

- Stufflebeam, D. L., W. J. Foley, W. J. Gephart, L. R. Hammond, H. O. Merriman & M. M. Provus 1971 Educational Evaluation and Decision-making in Education. Ithaca: Peacock Press
- Swain, M., S. Lapkin, N. Rowen & D. Hart 1990 The role of mother tongue literacy in third language learning. *Language, Culture and Curriculum* Vol 3 No 1 pp. 65-82
- Swales, J. 1989 Service English programme design and opportunity cost. in R.K. Johnson (Ed.) 1989: 79-90
- Swan, M. 1985 A critical look at the communicative approach 1 & 2. *English Language Teaching Journal* Vol 39 No 1 pp. 2-11; No 2 pp. 76-87
- Tawney, D. (Ed.) 1976 Curriculum Evaluation Today: Trends and Implications. London: Macmillan
- Ullman, R. 1990 Using complementary approaches to evaluate second language programmes. *Language, Culture and Curriculum* Vol 3 No 1 pp. 19-38
- Usher, R. & I. Bryant 1989 Adult Education as Theory, Practice and Research: the Captive Triangle. London: Routledge
- van de Ven, A. & A. Delbecq 1972 The nominal group as a research tool for exploring health studies. *American Journal of Public Health* Vol 62 No 4 pp. 337-342
- van Lier, L. 1988 The Classroom and the Language Learner. Harlow: Longman
- Wallace, M. J. 1991 Training Foreign Language Teachers. Cambridge: Cambridge University Press
- Wallace, M. 1998 Action Research for Language Teachers. Cambridge: Cambridge University Press
- Weick, K. 1976 Educational organisations as loosely-coupled systems. in *Administrative Science Quarterly* Vol 21 No 1 (Reprinted in Westoby (Ed.) 1988)
- Weir, C. & J. Roberts 1994 Evaluation in ELT. Oxford: Blackwell
- Weiss, C. 1986 Towards the future of stakeholder approaches in evaluation. In E. R. House (Ed.) *New Directions in Educational Evaluation*. pp. 186-200
- Wenden, A. 1998 Metacognitive knowledge and language learning. *Applied Linguistics* Vol 19 No 4 pp. 515-537
- Westoby, A. (Ed.) 1988 Cultures and Power in Educational Organisations. Milton Keynes: Open University Press
- White, R. 1988 The ELT Curriculum. Oxford: Blackwell

Bibliography

Widdowson, H. G. 1978 *Teaching Language as Communication*. Oxford: Oxford University Press

Widdowson, H. G. 1985 A Response to Swan. *English Language Teaching Journal* Vol 39 No 3 pp. 158-161

Wilcox, B. 1992 *Time-constrained Evaluation - A Practical Approach for LEAs and Schools*. London: Routledge

Wilkins, D. A. 1976 *Notional Syllabuses*. Oxford: Oxford University Press

Woods, D. 1996 *Teacher Cognition in Language Teaching*. Cambridge: Cambridge University Press

Woods, P. 1986 *Inside Schools: Ethnography in Educational Research*. London: Routledge & Kegan Paul

Wright, T. 1987 *Roles of Teachers and Learners*. Oxford: Oxford University Press

Yin, R. K. 1984 *Case Study Research: Design and Methods*. Beverly Hills: Sage

Young, R. 1992 *Critical Theory and Classroom Talk*. Clevedon: Multilingual Matters

Zeichner, K. M. 1993 Action Research: personal renewal and social reconstruction. *Educational Action Research* Vol 1 No 2 pp. 199-219

Data used for this study

CS1

<i>Nature of data-set</i>	<i>Volume</i>
Classroom notes (12X1.5hr sessions)	20,920 words
Interviews with Millie (3)	7,136 words
Materials (12X1.5hr sessions)	19 classroom activities

CS2

<i>Nature of data-set</i>	<i>Volume</i>
Classroom notes (12X3hr sessions)	54,310 words
Interviews with Anna (6)	15,509 words
Interview with May (1)	1,834 words
Interviews with students (23)	51,315 words
Transcript of small group discussion	2,153 words
Materials (12X3hr sessions)	39 classroom activities

Classroom Evaluation Case Study

Questionnaire on Evaluation for teachers of EAP programmes

Below is the questionnaire we discussed. Please fill in as a preparation for an interview. Especially valuable would be any events, stories which come to mind as you reflect on the questions/cues below. Continue on another sheet in necessary.

Thanks

Richard

1 *Professional biography*

1.1 Number of years teaching:

1.2 Evaluation in initial professional training:

1.3 Evaluation in teaching activity - key events:

1.4 INSET focussed on evaluation:

1.5 Evaluation policy induction:

1.6 Views on role of evaluation now:

2 *Describing evaluation practice*

2.1 Why?
(Policy; timing; classroom need)

2.2 What?
(Focus: Programme needs assessment - [goals]

Individual needs assessment - [learners; SLA frameworks]

Resource allotment - [time]

Appendix 2

Process or strategies for providing services to learners [activities; feedback to ss; materials; environment]

Outcomes of instruction - [tests; assessment activities])

(Sanders 1992:5-6)

2.3 How?
(Method)

3 *Evaluating evaluation practice*

3.1 Is it useful?
(Users; dissemination; action)

3.2 Is it practicable?
(Good use of class time; your time)

3.3 Does it do any damage?

Appendix 2

(Ss motivation; Class atmosphere; Other teachers)

3.4 Are findings accurate?

(Triangulation; Data-conclusion relationship)

*(after Joint Committee on Standards for Educational
Evaluation/Stufflebeam 1990)*

Interview schedule 1 - Teacher

Did the activities work as planned?

(cue each of the activities in turn; elicit history of tasks and texts; explore rationale for choosing activities/tasks; elicit views on how well they worked)

What did the activities tell you about the needs/wants of the group?

(cue with certain individual ss; specific issues/moments in the session; explore differences with groups in the past)

What plans for the next/future sessions?

(elicit outline for next session; explore rationale for tasks and activities, especially changes)

Did my presence have any effect in the session?

(on what T. did; on how T. planned; on ss in session)

Interview schedule 2 - Students

How did the activities work for you? Why appropriate / inappropriate, etc.?

(Elicit perceptions of experience; explore understanding of learning rationale; explore perception of needs)

Have you done activities like this before?

(Explore previous experience of activities; how s. relates activity in question to past experiences in classrooms)

Why do you think the teacher on this course does this kind of activity?

(Explore perceptions of group needs; perceptions of teacher's view of group; extent to which they feel the T. is attending to their needs and has got it right. Also ss perceptions of activity as typical of institution, or of this course only)

What kind of follow-up would be useful for you now?

(cue with homework from class; independent study in Library/Language Centre; work on other courses)

Have these interviews changed the way you see the course?

(Explore perceptions of me; of dealing with questions about learning)

Interview schedule 3 - Manager

What does the role of coordinator involve?

(Explore admin/bureaucratic and academic roles; who coordinator talks about courses to)

What are the main needs of students doing these courses in this institution?

(Explore different module profiles; different student profiles; how these have built up through practice)

What kind of activities or materials work well?

(Explore categories and sources of materials; reasons for preferences)

Are these shared by all the teachers involved?

(Explore shared and differing views of teaching activities and materials)

How important in the use of SA resources?

(Explore views; reasons; practices)

Do you get evaluation findings/reports from all the teachers involved?

(Explore practice, benefits, problems)

What do you do with these?

Explore what happens for curriculum development - discussion with team; and for accountability - reports to others)

Appendix 4

CS2 - Interviews with teacher, coordinator and students

Wk →	1 (1) 12/2	2(4) 19/2	3(3) 26/2	4(1) 5/3	5(1) 12/3	6(4) 19/3	7(5) 9/4	8(2) 16/4	9(5) 23/4	10(1) 30/4	11(1) 7/5	12(2) 14/5
Int/ee↓												
Anna	17/2	25/2	4/3			26/3			28/4			14/5
May									22/4			
Annie									29/4			
Arnie							9/4					
Dave						19/3			23/4			
Della							10/4					
Ina						20/3						
Isa		24/2										
Isa & Jamie				5/3								
Jamie											7/5	
Joe							10/4					
Laure		20/2	27/2			20/3	10/4		24/4			
Laure & Ina												14/5
Mat		19/2	26/2									
Niki								16/4				
Sona & Mari								17/4				
Sue & Sara					12/3					30/4		
Tigi							14/4					
TOTAL	(1) 12/2	(4) 19/2	(3) 26/2	(1) 5/3	(1) 12/3	(4) 19/3	(5) 9/4	(2) 16/4	(5) 23/4	(1) 30/4	(1) 7/5	(2) 14/5

CS2 - Classroom materials

Key**Activity code**

The dates of sessions (12 February - 14 May) and sequence of use of each activity

Materials description

Brief characterisation of what materials involve

Materials source

Where the materials originate from

Time (minutes)

The time in minutes (to nearest 5) spent on the materials in the classroom. (The total for each session should be 180 minutes. However this includes a break (approx. 30 minutes) and delays at the start of the session while waiting for students.

Pedagogic focus

Two broad categories are used here:

Meta-EAP = a focus on knowledge and skills which relate specifically to developing EAP skills - where learning might be considered directly transferable to the demands of studying other subjects through English

Within this category a binary distinction (after Wilkins 1976) is made between an **analytic** focus, where there is attention to EAP elements of a single complex discourse, and a **synthetic** focus, where there is attention to discreet linguistic elements of academic discourse.

Developing language skills = a focus on language use and form more generally, where learning might be considered only indirectly transferable (i.e. through improved English language skills) to the demands of studying other subjects

This category represents activities where the primary focus is language practice. The following sub-glosses are included to illustrate the main skills involved. There is no L for listening: all the listening practice involved was in interactional activities, i.e. there were no specific listening skills development activities.

S = speaking

R = reading

W = writing

V = vocabulary

CS2 - Classroom materials

Activity code	Materials description	Materials source	Time (minutes)	Pedagogic focus M = meta-EAP D = developing language skills
12/2/1	Module Guide - Procedural information on Content, Homework, Evaluation and Assessment, and a Reading / Viewing list of materials for SA use	Prepared by group of EAP teachers.	35	M - Anal
12/2/2	Guide to referencing - 2 pages	Prepared by T	20	M - Syn
12/2/3	Introductory Essay - Me - 1 page (mainly for ss to write)	Prepared by T	35	D - W
12/2/4	H/o on Discourse markers - Listing - 1 page	Prepared by T	15	M - Syn
12/2/5	Reading from the Economist on Computer Software - 2.5 pages	Downloaded from Internet by T	45	D - RSV
19/2/1	H/o on Writing types (with Purpose and Example)(a la Hedge) 1 page	Prepared by T	25	M - Syn
19/2/2	H/o for needs analysis / self evaluation - List difficulties in writing - (mainly for ss to write)	Prepared by T	15	M - Anal
19/2/3	H/o on Discourse markers - adverbials of Addition - 1 page	Prepared by T	15	M - Syn
19/2/4	Reading from the Economist on Gender issues in education and at work - 4.5 pages	Downloaded from Internet by T	65	D - RSV
19/2/5	Writing task based on 19/2/4 - 1 page - (mainly for ss to write.	Prepared by T	15	D - W
26/2/1	H/o on types of written argument - 1 page	Adapted by T from published source	35	M - Syn
26/2/2	H/o on Discourse markers - adverbials of Transition, Summation, and Apposition - 1 page	Prepared by T	20	M - Syn
26/2/3	Reading from the Sunday Times on Ethnic Nationalism - 5 pages	Downloaded from Internet by T	70	D - RSV
26/2/4	Writing task on 26/2/3 - 1 page (mainly for ss to write)	Prepared by T	15	D - W
5/3/1	Reading from research Report on U/g writing - 3 pages	Adapted by T from published source	65	M - Anal
5/3/2	H/o on Discourse markers - adverbials of Result - 1 page	Prepared by T	10	M - Syn
5/3/3	Interactive activity: A personal Language History - 2 pages	Prepared by T	65	D - S
12/3/1	H/o on Discourse markers - adverbials of Replacement and Contrast - 1 page	Prepared by T	15	M - Syn

Appendix 5

12/3/2	Readings from subject areas: from New Scientist on the extinction of languages (photocopy with pictures) - 4 pages; from published source on Creative Accounting (photocopy); - 7 pages; from published source on Freud (scanned) - 5 pages; from Sunday Times on F. R. Leavis and English Literature (from Internet) - 3 pages	Compiled by T from various published sources	45	D - RSW
12/3/3	Evaluation discussion	Prepared by T	60	
12/3/4	Evaluation - writing task	Prepared by Researcher	15	
19/3/1	Guide on ways of showing and avoiding commitment - 1 page	Adapted by T from published source	35	M - Syn
19/3/2	Reading form The Economist on Universities - 5 pages	Downloaded from Internet	65	D - RSV
19/3/3	Guide and tasks on Referencing in Academic writing	Adapted from published source	45	M - Syn
9/4/1	Guide to Interpreting assignment and examination questions - 3 pages	Adapted from published source	15	M - Syn
9/4/2	Guide and tasks to using quotations in academic writing - 3 pages	Adapted from published source	50	M - Syn
9/4/3	Readings from The Guardian on arguments for and against closer European Union - 1 page for, and 1 page against	Photocopied from The Guardian	60	D - RSV
16/4/1	Guide on Techniques for making a Presentation 4 pages	Adapted from published source	30	M - Syn
16/4/2	Reading from Research Report on U/g writing (Part 2) - 3 pages	Adapted from published source	30	M - Anal
16/4/3	Reading from newspaper on global geopolitics - the fate of the West - 1 page	Photocopied from newspaper	75	D - RSV
23/4/1	Reading and task on Academic success - 2 pages	Adapted from published source	40	M - Anal
23/4/2	Reading from The Economist on private policing- 4 pages	Downloaded from the Internet	95	D - RSV
30/4/1	Guide to, and series of tasks on Writing Styles	Adapted from published source	60	M - Anal
30/4/2	Reading from The Economist on Terrorism	Downloaded from the Internet	70	D - RSV
7/5/1	Ss' Presentations	Students' work	70	D - S
7/5/2	Reading and Task from the Sunday Times on Violence against University students - 3 pages	Downloaded from the Internet	75	D - RSV
14/5/1	Ss' Presentations	Students' work	70	D - S
14/5/2	Reading and task from the Economist on Computer Software - 2.5 pages as 12/2/2 with task	Downloaded from Internet by T	75	D - RSV
14/5/3	H/o on Cohesive Markers- 1 page	Adapted from published source	- (handed out)	M - Syn

Evaluation Policy

from University Handbook of Policies and Procedures

Quality Management Section

STAKEHOLDER EVALUATION

I. GENERAL PRINCIPLES

- A. Stakeholder evaluation is undertaken primarily for two purposes: accountability and development. The University is committed to use evaluation to inform decision making and to aid the development of an effective teaching/learning environment.
- B. Evaluation methods used must demonstrate a balance of quantitative and qualitative approaches and seek both internal and external evaluative data.
- C. Evaluation methods must use appropriate criteria and systematically collect information so that the quality and effectiveness of modules and programmes can be assessed.
- D. Stakeholders should feel that they own evaluation through active participation in the process and by seeing that their evaluations are being used to make positive changes.
- E. Evaluation methods must be user friendly and the resulting information should be easy to collate and should be communicated effectively.
- F. In order that the evaluation process is implemented effectively staff need training. Conversely, involvement in evaluation leads to staff development.
- G. Stakeholder evaluation should not be confused with staff appraisal, for which entirely different systems must be set up.

II. THE STAKEHOLDERS

- A. The stakeholders of the university's programmes are considered to be the following:

Students - full and part-time
Parents and partners
Employers as sponsors of students
Employers as receivers of students
Professional Bodies
Awarding, examining and validating bodies
Funding Bodies (Public sector)
Staff - teaching and support
Franchisees and overseas links
Corporate Services users

III. WHO AND WHAT WILL BE EVALUATED?

- A. In principle, all aspects of the activities of the University should be evaluated. The activities can be divided into three main areas, educational provision, support services, and the University management. Below is shown a breakdown of the aspects of the University's activities which should be evaluated. This is not intended to be an exhaustive list.

Teaching/learning quality - by students, peers and self
Teaching/learning delivery modes
Teaching/learning resources
Teaching/learning environment
Assessment results
Welfare services
Student social life
Employment opportunities
Value for money
Level of competence of students (initial and longer term)
University infrastructure and administration
Department management
Directorate

IV. METHODS OF EVALUATION

- A. Evaluation will be carried out in every dimension of the teaching/learning environment. Evaluators will be required to provide a rationale of the methods they have chosen and to demonstrate that a variety of methods have been used. Given the wide range of stakeholders it may not be possible to evaluate all of them every year, but it is expected that over a Four or Five year period each will be covered. This paper concentrates on the evaluation of the educational provision and the experiences of students. The evaluation of central services and management will need further consideration.
- B. Evaluation framework
- C. When a specific activity is being evaluated it may be helpful if the evaluators think through the process using the following framework.
- Aims of the evaluation
Data collection process
Data processing
Distribution of results
Actions to be taken
- D. Possible evaluation methods
- Questionnaires
Structured group feedback
Diaries
Peer review
Self-evaluation
Interviews
External Examiners' reports
Employer and placement feedback
- E. A pro-forma questionnaire for completion by students has been produced and this is available for use by programmes and modules.

V. FLOW OF EVALUATION INFORMATION

A. TEACHING/LEARNING

The responsibility for the evaluation of each module rests with the SOG which, in conjunction with the host School's Teaching Resource Manager, will determine the methods to be used. The results of the evaluation, including action to be taken, will be fed via the TRM to the School Board and, for information, to Programmes Managers.

The responsibility for the evaluation of each programme rests with the programme team which, in conjunction with the Programmes Manager, will determine the methods to be used. As part of this process, information may be collected from external sources, such as employers and professional bodies. Programme evaluation will concentrate on the structure of the programme and the entire student experience rather than the academic content, which is covered by the module evaluation. The results of the evaluation, including action to be taken, will be fed via the Programmes Manager to the Departmental Board.

The responsibility for the evaluation of Library Services, Corporate Services and Central Services rests with the Head of the Unit who will determine, in conjunction with the member of the Directorate to whom the unit reports, the evaluation methods to be used. The results of the evaluation, including action to be taken, will be fed through to the Directorate.

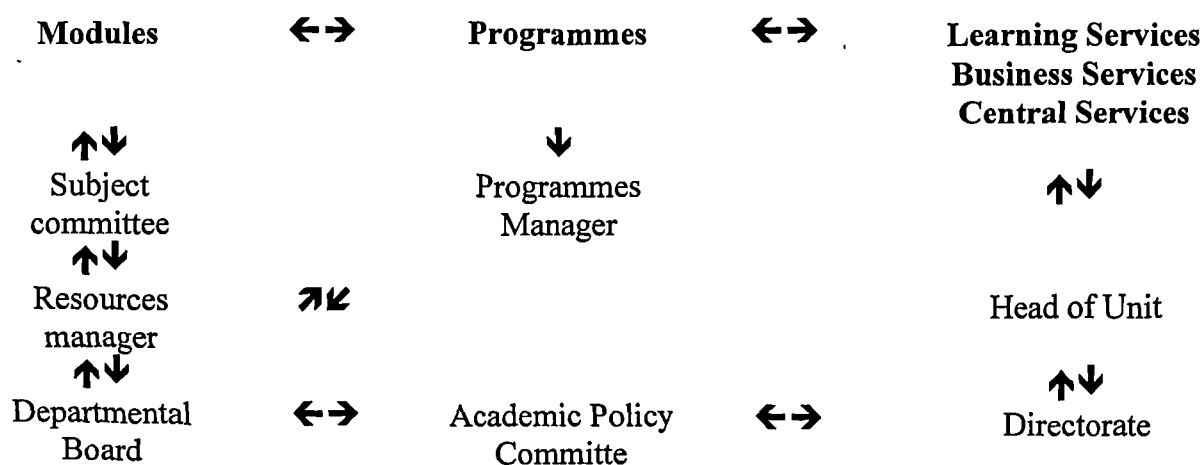
B. DEPARTMENTAL BOARDS

The School Board will monitor the results of the evaluation of programmes and modules and will produce an annual report to the Academic Policy Committee highlighting major problems and common issues and the action to be taken to remedy them.

C. FEEDBACK

As the purpose of evaluation is for development as well as accountability, it is crucial that feedback takes place, at the appropriate time, so that appropriate action can be taken immediately, or at least in time to affect the module or programme the next time it is delivered. This implies that evaluation will take place during, as well as at the end of, a module or programme. The diagram below indicates the feedback loops which take place at each stage.

Diagram of Information Flows



VI IMPLEMENTATION

A. INTRODUCTION

This approach tries to avoid being over prescriptive about the evaluation process so that Departments who already possess a high level of awareness of what a quality evaluation process involves will find themselves uncomfortably restricted by guidelines and suggestions that are too rigid or are interpreted in too inflexible a manner. On the other hand, the evaluation process may fail if insufficient support or guidance is not provided for those who are less experienced in the implementation of the evaluation process.

B. MANAGEMENT OF THE EVALUATION PROCESS

It would be helpful if the evaluation process could be centrally coordinated and for departments to have a central source of expertise to call upon to support their evaluations.

C. EXAMPLES OF GOOD EVALUATION PRACTICE

To support evaluation within the university, a number of inputs should be produced, eg, the development and coordination of examples of good evaluation practice within the institution. Sample evaluation procedures and frameworks for the different stages in the evaluation process (viz., planning, implementation, reporting and acting) should be made available. It is recommended that pilot evaluation activities/studies are encouraged in the institution such that a bank of evaluation styles are developed.

D. STAFF DEVELOPMENT

Successful implementation of evaluation across all of the University's activities requires a significant amount of staff development. It is proposed that SOG Co-ordinators, Programme Leaders, Programmes Managers, Teaching Resource Managers and interested staff should attend a course which will inform them of what the various evaluation techniques are and how to choose the appropriate ones for their purposes. In time all staff should be given the opportunity to attend such a course.

Departmental policy on evaluation interpretation and implementation strategy for institutional Evaluation Policy

English Language Programmes Group (ELPG) policy is that all modules should a) be evaluated after four weeks by means of a classroom discussion, using the nominal group technique, and b) use a final evaluation questionnaire at the end of the module.

The procedure is as follows:

1. Posing the questions

Two questions are written on the board, one dealing with positive aspects of the module, and the other dealing with the negative. Example questions might be:

What three aspects of the work in this class have helped you to improve your knowledge of _____ / did you most enjoy?

What three aspects of the work in this class have not helped you to improve your knowledge of _____ / did you least enjoy?

2. Silent nominations

Individuals are given, say, 5/10 minutes, to list their responses on a piece of paper/card. Complete silence is advisable.

3. Master lists

The teacher compiles two master lists on the BB/OHP taking only one positive item from each member of the class in rotation, until each student has nominated three items or the list is exhausted. The procedure is repeated for negative items. No editing is allowed and no evaluative comment (by the teacher) is made at this stage. It is helpful to number the items.

4. Item clarification

Each item is discussed by the whole class until each student is clear about what it means. If a student feels that their item is already covered by someone else, they may request that the item is withdrawn.

5. Final listing by students (Silent)

Students choose from the master list their three most important positive points and three most important negative points, this time in order of priority.

6. Teacher prepares final evaluation

By whatever means she / he prefers - one possibility is to award the first listed item three points, the second two points, etc. - the teacher prepares a list of the group's priorities, and

a) presents the list to the class for discussion in the next lesson. If required / necessary / possible, an agreed change in the course plan for teaching and learning can be made.

b) The list of class priorities, together with any changes of procedure or content agreed with the students, should be lodged with the ELPC coordinator and with other programme coordinators, as appropriate.

Writing task CS2:1

Essay Title : Me

Write a short essay about yourself. (We will give you about 40 minutes to do it). You are free to include whatever information you wish. However you might want to include information such as your name; your home country; the kind of education you have had there; the languages you speak; what made you come to England to study; what programme you are studying at TVU; what your future plans are what you hope to achieve by studying this module and so on. This will help us to form an idea of what your writing in English is like. This will assist us in planning the module in a way that will give you the maximum opportunity to work out the things you need to practise.

Two samples, typed up from copy of student manuscript*Arnie*

My name is Arnie and I come from Nice in the South of France. I've always been interested in languages rather than mathematics, so I followed the way of English and Italian from high school until now.

As the university only teaches you theory, I wanted to put into practice my knowledge in English coming abroad. I hesitated a long time between going to Italy or England because I like Italian too and I live very near to the border. Finally I preferred going to England because English is the most important language in the world. Then I chose [...] University because it was situated in London and so it was a good opportunity to discover the capital.

I miss Italian because I have only courses in English. The level of Italian is only for beginners, so I will have to sit my exams in June in France.

I wasn't sure that I could stay one year in a foreign country far from my home and family, but in fact it works out. Life in London is nice, my "English family" is very kind and I can meet a lot of Erasmus students and English people as well.

Sue

Personal dates [data]

As you can see from the top, my name is Sue. I'm 23 years old, female, from Germany, from a small city called Erlangen, which is in the south part of Germany called Bavaria. I'm studying psychology for a Diploma degree in Germany, and I'm nearly finished right now. My home university is called OHO - Friedrich Universität Bamberg and is not far away from Erlangen. I speak German and English language only.

Reasons to study in England

I think my main reason to come here was to improve my personal experiences and my English. because of the fact that I'm nearly finished with my psychology studies in Germany, I wanted to do something different from just running through university and starting to work afterwards. In addition I have to wait for a new professor of Human Resource Management, my minor subject in Germany, and because of this, I decided to use my time to study Psychology in England to get a different view of Psychology. I chose London, because I've been here four years ago for holidays, and I really loved the town, the flair, the people with their different cultures, religions and languages, which I often miss in Germany.

Future plans

I will leave this university and London in the middle of May to study the German Summer semester at home as well, because the German government changed the conditions of our sponsorship which will cause financial problems for me in the future. After finishing my studies I hope to work as an occupational psychologist in personnel, for example, Human Resource Management, or Human Resource Development.

Achievements by studying this module

I hope to improve my writing, reading, understanding, discussing and speaking in English language.

Writing Task CS2:5

Evaluation - Reflections on Expectations and Experience

Please write a summary of your expectations and experience of this module - 200-500 words. Include:

- what you expected in this module;
- how your experience of the module compares with expectations;
- what you need to concentrate on now.

Sao

The name of the module is English for Academic Purposes. It makes me interested in this module. It is a new subject that I never learn it before. The module is rather difficult for me to learn at the first time. It is sad that I didn't learn at the first [module in] semester 1. I have started to learn in the second semester. I try to understand the purposes of academic writing. One thing I know, that is use for higher study at university. I have expected in this module. There are many expected that I want in this module. Improving my writing essay in correct way. I can use transition, summation, and opposition in the sentence. I can learn style of writing essay and which style is suit me. When I learn general English at school, it doesn't teach about writing or style of writing. the teacher gives a topic and do homework. It was writing letters and descriptions. When the lecturer gave texts about style of writing in class for the students read the texts. I read through the whole text to get an idea of the subject matter, the overall meaning and the vocabulary. I find it's difficult. What does about the text? Who wrote it? Why? What does the point of view of the writer? What do you think about the text? I can improve my reading than before. I can have group discussion with the other friends to exchange an idea. Sometimes my friends help me to summarise about the text. It is very helpful. I think I should read more, practise writing and practise reading. I hope it is not late to do it. As comparison, I find the academic writing need more information in your own idea and I should [have] analysis in my own writing.

At the end, I should read lots of books and try to catch up the point of view of the writer and summary of the text that I read. Moreover I should apply or interpret idea and get idea in the writing. I know it's difficult for me to do that. I think I need more reading. The more you read you more get idea.

Laure

I chose this module as I really want to improve my English (the way of speaking and writing in English). In this module I thought that we will do some writing and we are doing it - I'm sure I will improve my writing as every week we have to give you a short task about a subject studied in class and that I find very useful if the piece of paper about the link words - then I can say that we learn how to structure an essay and even if at the beginning we always forget to use the link words, I am sure that within two weeks we will become used to structure our essays properly.

Moreover, I find it very useful the work group as we can exchange our points of view and help each other to understand better the text if there are some words that we don't know. Even if, it isn't really easy for us to speak in front of the class it is useful for us as it is a good practice to have more confidence in ourselves and for a future interview for a job. In fact, I say that as if we are never used to speak in front of an audience we are very embarrassed when we have to do it so it is better to have practice now and it seems also to be a good training for our next oral presentation.

But, on another side, I find the idea to do more vocabulary really interesting to improve our vocabulary - I also want to say that I like the work discussions as each student has an opportunity to exchange his point of view and to debate on it.

As a conclusion, I will say that I have to concentrate myself on the structuration of an essay and also, when I am speaking, I have to try to avoid important grammar mistakes. My aim is to acquire a perfect grammar and speaking in English.